

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 14:04:54 ; Search time 2174 Seconds  
(without alignments)  
9451.049 Million cell updates/sec

Title: US-09-912-628-4  
Perfect score: 706  
Sequence: 1 ggcacgagctctctctctg.....gtaaaaaaaaaaaaaaaa 706

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_ram.\*
- 37: em.htg\_vrt.\*
- 38: em.ey.\*
- 39: em.htgo\_hum.\*
- 40: em.htgo\_mus.\*
- 41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	250.4	35.5	695	6	AX147521	Sequence
2	250.4	35.5	868	6	AX147520	Sequence
3	250.4	35.5	1302	6	AX147519	Sequence
4	249.2	35.3	1366	6	AX116612	Sequence
5	248.4	35.2	1565	9	AK098566	Homo sapi
6	247.2	35.0	1361	9	HSTHRINH	222658 H. sapiens t
7	246.8	35.0	1324	6	AX147509	Sequence
8	246.8	35.0	1349	6	AX147468	Sequence
9	246.8	35.0	1410	6	AX147508	Sequence
10	246.8	35.0	1465	6	AX147508	Sequence
11	246.8	35.0	1465	6	AX147508	Sequence
12	246.8	35.0	1523	9	BC001394	cytoplasmic
13	246.8	35.0	2036	9	AK094983	Homo sapi
14	246.8	35.0	3271	9	AK057138	Homo sapi
15	245.2	34.7	2759	9	AK027149	Homo sapi
16	244	34.6	1327	4	D55670	Homo sapi
17	237.8	33.7	1317	10	MMU25844	Bos taurus
18	236.6	33.5	1532	10	BC006766	Mus muscu
19	225.4	31.9	1859	10	U96707	Mus muscu
20	218	30.9	1265	10	AF425084	Mus muscu
21	214.4	30.4	311	6	AX186945	Sequence
22	214.4	30.4	1319	9	BC034528	Homo sapi
23	214.4	30.4	1325	9	HUMCAP2A	Homo sapi
24	214.4	30.4	1425	6	AR005195	Sequence
25	214.4	30.4	1425	6	AR005195	Sequence
26	214.4	30.4	1425	6	AR005195	Sequence
27	212.4	30.1	922	9	HSU92984	Sequence 1
28	206	29.2	186351	2	AP001503	Homo sapien
29	206	29.2	191395	9	AC009802	Homo sapi
30	206	29.2	191793	2	AP001897	Homo sapi
31	206	29.2	198231	2	AC067875	Homo sapi
32	206	29.2	201734	2	AC090224	Homo sapi
33	196.8	27.9	287	6	AX185519	Sequence
34	179.2	25.4	1626	9	HSU71364	Human serin
35	178.8	25.3	1413	10	U96703	Mus musculus
36	178.8	25.3	1779	10	BC025121	Mus muscu
37	177.4	25.1	1385	9	HUMCAP3A	Homo sapien
38	177.4	25.1	1393	6	AR005196	Sequence
39	177.4	25.1	1393	6	AR005196	Sequence
40	177.4	25.1	1393	6	AR005196	Sequence
41	177.4	25.1	2792	9	AB060690	Sequence 3
42	177.4	25.1	4130	9	BC002538	Homo sapi
43	167.4	23.7	1819	10	U96700	Mus muscu
44	167.4	23.7	2283	10	BC029900	Mus muscu
45	152.4	21.6	1931	10	U96705	Mus muscu

ALIGNMENTS

RESULT 1	AX147521	Sequence 75 from Patent WO0136632.	695 bp	DNA	linear	PAT 08-JUN-2001
LOCUS	AX147521	Sequence 75 from Patent WO0136632.	695 bp	DNA	linear	PAT 08-JUN-2001
DEFINITION	AX147521	Sequence 75 from Patent WO0136632.	695 bp	DNA	linear	PAT 08-JUN-2001
ACCESSION	AX147521	Sequence 75 from Patent WO0136632.	695 bp	DNA	linear	PAT 08-JUN-2001
VERSION	AX147521.1	GI:14346678	695 bp	DNA	linear	PAT 08-JUN-2001
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 695)					
AUTHORS	Levine, Z., David, A., Azar, I., Khosravi, R. and Bernstein, J.					
TITLE	Patent: WO 0136632-A 75 25-MAY-2001;					
JOURNAL	Patent: WO 0136632-A 75 25-MAY-2001;					

Db	CTTGGGCTCGCCATCATGGATGTTCTCGCAGAAGCAAAATGGCACCTTTGCTTAAACCTT	119
Qy	TTGAAAAAGCTAGGGGAAAAAACAACCAACTTATTTTTTTC-CCCATGAGCATATCA	155
Db	TTGAAAAACGCTGGGTAAAGACAACTCGAAGAAATGTTTTTCTCACCATGAGCATGTC	179
Qy	TCAGCCTTGGCCATGGTTTTTCATGGGGCAAGGAAAAACATGCGAGCTCAGATGCTCAG	215
Db	TGTGCCCTGGCCATGGTCTACATGGGGCAAGGAAAAACCGCTGCACAGATGGCCAG	239
Qy	GCATTTGTTTGTAGTAAATCGAGGTGAAGATGAGATATTTCATCGAGGTTTTTCAGTCA	275
Db	ATACTTTCTTCAATAAAGTGGCGGT---GGTGGAGACATCCACAGGGCTTCCAGTCT	296
Qy	CTTCTTTGTCATTAACAGAACTGACACTGAATATGCTTTAGAACTGCCAACGGGCTC	335
Db	CTTCTCACCGAAGTGAACAGACTGGCAGCGAGTCTGCTTAGGATGGCCAAACAGGCTC	356
Qy	TTTGGAGAAAAGTCTTATGATTTCTCCACAGGTTTTTACAGATTCCTGTGGCAATCTAC	395
Db	TTTGGGAAAAGTCTTGTGATTTCTCTCATCTTTTAGAGATTCCTGCCAAAATTTCTAC	416
Qy	CAAGCAACGATAAAACAGCTAGACTTTGTGAATGATACAGAGAAGTCCCAACACGTGA	455
Db	CAAGCAGAGATGGAGGAGCTTGACTTTTATCAGCGCGGTAGAGAGTCCAGAAAACACATA	476
Qy	AACCTCTCTGGGTGCTGATATAAACTAAAG	483
Db	AACACCTGGGTAGCTGAAAAGACAGAG	504
RESULT 3		
AX147519	AX147519	1302 bp DNA linear PAT 08-JUN-2001
LOCUS	Sequence 73 from Patent WO0136632.	
DEFINITION	AX147519	
ACCESSION	AX147519.1	GI:14346676
VERSION		
KEYWORDS	human.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 1302)	
TITLE	Levine, Z., David, A., Azar, I., Khosravi, R. and Bernstein, J.	
JOURNAL	Patent: WO 0136632-A 73 25-MAY-2001;	
FEATURES	CompuGen Ltd. (IL)	
source	Location/Qualifiers	
	1. 1302	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
BASE COUNT	341 a 305 c 345 g 311 t	
ORIGIN		
Query Match	35.5%; Score 250.4; DB 6; Length 1302;	
Best Local Similarity	75.4%; Pred. No. 8.4e-61;	
Matches	338; Conservative 0; Mismatches 106; Indels 4; Gaps 2;	
Qy	CTTGGGCTCGCCATCATGGATGTTCTCGCAGAAGCAAAATGGCACCTTTGCTTAAACCTT	96
Db	CTTGGGCTCGCCATCATGGATGTTCTCGCAGAAGCAAAATGGCACCTTTGCTTAAACCTT	119
Qy	TTGAAAAAGCTAGGGGAAAAAACAACCAACTTATTTTTTTC-CCCATGAGCATATCA	155
Db	TTGAAAAACGCTGGGTAAAGACAACTCGAAGAAATGTTTTTCTCACCATGAGCATGTC	179
Qy	TCAGCCTTGGCCATGGTTTTTCATGGGGCAAGGAAAAACATGCGAGCTCAGATGCTCAG	215
Db	TGTGCCCTGGCCATGGTCTACATGGGGCAAGGAAAAACCGCTGCACAGATGGCCAG	239
Qy	GCATTTGTTTGTAGTAAATCGAGGTGAAGATGAGATATTTCATCGAGGTTTTTCAGTCA	275
Db	ATACTTTCTTCAATAAAGTGGCGGT---GGTGGAGACATCCACAGGGCTTCCAGTCT	296
Qy	CTTCTTTGTCATTAACAGAACTGACACTGAATATGCTTTAGAACTGCCAACGGGCTC	335
Db	CTTCTCACCGAAGTGAACAGACTGGCAGCGAGTCTGCTTAGGATGGCCAAACAGGCTC	356
Qy	TTTGGAGAAAAGTCTTATGATTTCTCCACAGGTTTTTACAGATTCCTGTGGCAATCTAC	395
Db	TTTGGGAAAAGTCTTGTGATTTCTCTCATCTTTTAGAGATTCCTGCCAAAATTTCTAC	416
Qy	CAAGCAACGATAAAACAGCTAGACTTTGTGAATGATACAGAGAAGTCCCAACACGTGA	455
Db	CAAGCAGAGATGGAGGAGCTTGACTTTTATCAGCGCGGTAGAGAGTCCAGAAAACACATA	476
Qy	AACCTCTCTGGGTGCTGATATAAACTAAAG	483
Db	AACACCTGGGTAGCTGAAAAGACAGAG	504
RESULT 2		
AX147520	AX147520	868 bp DNA linear PAT 08-JUN-2001
LOCUS	Sequence 74 from Patent WO0136632.	
DEFINITION	AX147520	
ACCESSION	AX147520.1	GI:14346677
VERSION		
KEYWORDS	human.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 868)	
TITLE	Levine, Z., David, A., Azar, I., Khosravi, R. and Bernstein, J.	
JOURNAL	Patent: WO 0136632-A 74 25-MAY-2001;	
FEATURES	CompuGen Ltd. (IL)	
source	Location/Qualifiers	
	1. 868	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
BASE COUNT	214 a 212 c 227 g 215 t	
ORIGIN		
Query Match	35.5%; Score 250.4; DB 6; Length 868;	
Best Local Similarity	75.4%; Pred. No. 8.3e-61;	
Matches	338; Conservative 0; Mismatches 106; Indels 4; Gaps 2;	
Qy	CTTGGGCTCGCCATCATGGATGTTCTCGCAGAAGCAAAATGGCACCTTTGCTTAAACCTT	96

Db 240 ATACTTTCTTCAATAAAGTGGCGGT---GGTGGAGACATCCACAGGCTTCCAGTCT 296  
Qy 276 CTTCTTGTGCAATTAACAGAACTGACACTGAATATGTGCTTAGAACTGCCACGGGCTC 335  
Db 297 CTTCTACCGGAAGTGAACAGACTGGCGCAGTACTTGTCTTAGATGGCCACAGGCTC 356  
Qy 336 TTTGGAGAAAGTCTTATGATTTCTCTCACAGGTTTACAGATTCTGTGGCAAAATTTCTAC 395  
Db 357 TTTGGGAAAGTCTTGTGATTTCTCTCATCTTTTAGAGATTCTTCCGCAAAATTTCTAC 416  
Qy 396 CAAGCAACGATATAACAGCTAGACTTGTGAATGATACAGAGAGTCCACACAGCTGTA 455  
Db 417 CAAGCAGATGAGAGGCTTGACTTTATCAGCGCGTAGAGAAGTCCAGAAACACATA 476  
Qy 456 AACTCTGGTGTCTGATATAAACTAAAG 483  
Db 477 AACACCTGGGTAGCTGAAGAAGACAGAAG 504

RESULTS 4  
ARL16612  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
BASE COUNT 360 a 332 c 355 g 317 t 2 others  
ORIGIN

Query Match 35.3%; Score 249.2; DB 6; Length 1366;  
Best Local Similarity 73.4%; Pred. No. 1.9e-60;  
Matches 345; Conservative 1; Mismatches 120; Indels 4; Gaps 2;

Qy 15 TCTGGGAGCTCGCGGAGAACTGGGCTCACCTCATGATGCTCTATCAGAAAGCAAA 74  
Db 40 TTCTGCTGCTCGCTCCCGCTCTGGAGTCTGCCATCATGGATGTTCTCGCAGAGCAAA 99  
Qy 75 TGGCATTGTGCAATTAACCTTTTCAAAAAGCTAGGGAAACAACTCAAACTTATT 134  
Db 100 TGGCACCCTTTGCTTTAAACCTTTTGAACAACTGGGTAAAGCAACTCGAAGATGTGT 159  
Qy 135 TTTTTC-CCCATGACATATCATAGCTTGGCCATGTTTTCATGGGGGCAAAAGGAA 193  
Db 160 TTTTCACCCATGAGCATGTCTGTGCGCTGGCCATGCTTACATGGGGGCAAAAGGAA 219  
Qy 194 CACTGCAGTCTAGATGTCTCAGGCACCTTTGTTTGTAGTAAATCGAGGTGAAGTGA 253  
Db 220 CACCGCTGCACAGATGGCCAGATACTTCTTTCAATAAAGTGGCGGTG---GTGGAGA 276  
Qy 284 TATTATCGAGTTTTCAGTCACTTCTTTGTGCAATTAACAGAACTGACACTGAATATGT 313  
Db 277 CATCCACAGGCTTCCAGTCTCTTCTCACCAGTGAACAGACTGGCAGCGAGTACTT 336  
Qy 314 GCTTAGAACTGCCACGGCTTTTGGAGAAAGTCTTATGATTTCTTCACAGGTTTAC 373  
Db 337 GCTTAGGATGCCCAACAGGCTTTTGGGAAAGTCTTTGTGATTTCTCTCATCTTTAG 396  
Qy 374 AGATTCTGTGGCAAAATTTCTACCAAGCAACGATATAAAGCAGCTAGACTTTGTGATGATAC 433  
Db 397 AGATTCTGTGCCAAAATTTCTACCAACBAGATGGAGAGCTTGACTTTATCAGCGCGT 456  
Qy 434 AGAGAGTCCCAACAGCTGTAAACTCTGTGGTGTCTGATATAAACTAAAG 483

Db 457 AGAGAAGTCCAGAAAACACATAAACACCTGGTAGCTGAAAAAGACAGAAG 506

RESULTS 5  
AK098566  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
BASE COUNT 390 a 381 c 428 g 366 t  
ORIGIN

Query Match 35.2%; Score 248.4; DB 9; Length 1565;  
Best Local Similarity 76.0%; Pred. No. 3.2e-60;  
Matches 333; Conservative 0; Mismatches 101; Indels 4; Gaps 2;

Qy 47 CCGTCATGATGCTCTATCAGAAAGCAATGGCACAATTGCAATTAACCTTTTGAAGAAGC 106  
Db 296 CCATCATGGATGTTCTCGCAGAAAGCAATGGCACCTTTGCTTTAAACCTTTTGAAGAAGC 355  
Qy 107 TAGGGCAAAACAACTCAAACTTATTTTTC-CCCATGAGCATATCATCAGCCTTG 165  
Db 356 TGGGTAAGACAACACTCGAAGATGTGTTTTCTCACCCATGAGCATGTCTGTGCCCTGG 415  
Qy 166 CCATGTTTTCATGGGGGCAAAAGGAAACACTGCGAGCTCAGATGTCTCAGGCACTTTGTT 225  
Db 416 CCATGCTTACATGGGGGCAAAAGGAAACACCGTGCACAGATGGCCAGATCTTTCTT 475  
Qy 226 TTAGTAAATCGAGGTGAAGATGGAGATTTATCATGAGGTTTTCAGTCACTTTCTGTTG 285  
Db 476 TCAATAAAGTGGCGGT---GGTGGAGACATCCACAGGCTTCCAGTCTCTTCTCACCG 532  
Qy 286 CAATTAAACAGAACTGACACTGAATATGTCTTTAGAACTGCCAACCGGCTCTTTGGAGAA 345  
Db 533 AAGTGACAGAACTGGCAGCGAGTACTTCTTAGATGGCCAAAGGCTCTTTTGGGAAA 592

AK098566  
Homo sapiens cDNA FLJ25700 fis, clone TST04708, highly similar to  
PLACENTAL THROMBIN INHIBITOR.  
AK098566  
AK098566.1 GI:21758608  
oligo capping; fis (full insert sequence).  
Homo sapiens testis cDNA to mRNA, clone\_lib:TST clone:TST04708.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,  
Arita, M., Mueashino, K., Yuuki, H., Hara, H., Suzuki, Y., Hata, H.,  
Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T.,  
Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A.,  
Kawakami, B., Nagai, K., Isogai, T. and Sugano, S.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 1565)  
Sugano, S. and Suzuki, Y.  
Direct Submission  
Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,  
University of Tokyo, Laboratory of Genome Structure, Human Genome  
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan  
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,  
Fax:81-3-5449-5416)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction and 5'-end one pass sequencing: Institute of Medical  
Science, University of Tokyo, Laboratory of Genome Structure, Human  
Genome Center; 3'-end one pass sequencing: RAB; clone selection for  
full insert sequencing: RAB and Helix Research Institute.  
Location/Qualifiers  
1..1565  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="TST04708"  
/csize="testis"  
/clone\_lib="TST"  
/notes="Cloning vector: pME18SFL3"

Wed May 28 14:44:07 2003

us-09-912-628-4.rge

QY 346 AGTCTTATGATTTCTCTCAGGTTTACAGATTCCTGTGGCAATCTTACCAAGACGA 405  
 Db 593 AGCTTTGATTTCTCTCATCTTTTAGAGATTCCTGCGCAAAATTTTACCAAGCAGAGA 652  
 QY 406 TAAACAGCTAGACTTTGTGATGATACAGAGAGTCCACACACGCTGTAACCTCTCTGGG 465  
 Db 653 TGGAGGAGCTTACCTTTATACAGCGCTAGAGAGTCCAGAAAACACATTAACACCTGGG 712  
 QY 466 TTGCTGATAAACTAAAG 483  
 Db 713 TAGCTGAAGACAGAG 730

RESULT 6  
 HSTRINH 1361 bp mRNA linear PRI 13-JUL-1994  
 LOCUS H.sapiens thrombin inhibitor mRNA.  
 DEFINITION Z22658  
 ACCESSION Z22658.1 GI:297411  
 VERSION thrombin inhibitor.  
 KEYWORDS Homo sapiens.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1361)  
 AUTHORS Coughlin, P., Sun, J., Cerruti, L., Salem, H.H. and Bird, P.  
 TITLE Cloning and molecular characterization of a human intracellular  
 serine proteinase inhibitor  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90 (20), 9417-9421 (1993)  
 MEDLINE 94022386  
 PUBMED 8415716  
 REFERENCE 2 (bases 1 to 1361)  
 AUTHORS Steinle, A.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-APR-1993) Steinle A., University of Munich, Institute  
 of Immunology, Goethestrasse 31, W-8000 MUENCHEN 2, GERMANY  
 FEATURES  
 source  
 /organism="Homo sapiens"  
 /macronuclear  
 /db\_xref="taxon:9606"  
 /clone="PTI/P"  
 /tissue\_type="placenta"  
 /clone\_lib="cDNA in lambda gt11 (clontech)"  
 75..3205  
 /functions="serine proteinase inhibitor"  
 /citations={1}  
 /codon\_start=1  
 /evidence=experimental  
 /label=PTI  
 /product="thrombin inhibitor"  
 /protein\_id="CA880373.1"  
 /db\_xref="GI:297412"  
 /db\_xref="SWISS-PROT:P35237"  
 /translation="MDVLAEGNTFALNLLKLTGKNSKNVFPSPMSMCAAMVYMG  
 AKGNTAQAQILSNFKSGGGDIHQGFQSLLEVNKTGTQLLRVANLFGKSCDF  
 LSSPRDSQFQOAEMLDFISAVEKSRKHINTVAEKEGKIAELLSFGSDPLTR  
 LVLVNAVYFGNWDGQDFKENTRLFKVSKNEKFPVMMFKOSTFEKTYIGEIFQI  
 LVLPLYGKELNMLIMLPDETDLRTVEKLTVEKFWENTLDMDEEVEVSLPRFL  
 EESYDMESVLRNLGMIDAFELGKADFSQMSQTLDSLSKVHKSFEVNEEGTEAAAAAT  
 AAINMMFCARFVPRFCADHPFLFIQHRKNTGILFCGRFSSP"

polya\_signal 353 a 328 c 319 t  
 BASE COUNT 353 a 328 c 361 g 319 t  
 ORIGIN  
 Query Match 35.0%; Score 247.2; DB 9; Length 1361;  
 Best Local Similarity 75.0%; Pred. No. 7e-60;  
 Matches 336; Conservative 0; Mismatches 108; Indels 4; Gaps 2;  
 QY 37 CTGGGGCTCAGCGTCATGATGCTCTATCAGAGCAAAATGGCATTGTCATTAACCTT 96  
 Db 60 CTTGGGCTGCATCATGATGTTCTCGCAGAGGCAAAATGGCACCTTTGCTTAAACCTT 119

QY 97 TTGAAAAAGCTAGGGGAAAAAACAACCTCAAAACAACCTTATTTTTTTC-CCCATGAGCATATCA 155  
 Db 120 TTGAAAAAGCTAGGGTAAAGCAACTCGAAGATGTGTTTTTCTCACCCTAGCATGTCC 179  
 QY 156 TCAGCCCTTGGCCATGTTTTCATGGGGCAAGGGAACACTGTCAGCTCAGATGTCTCAG 215  
 Db 180 TGTGCCCTGGCCATGTTTCTACATGGGGCAAGGGAACACCCGCTGCACAGATGGCCAG 239  
 QY 216 GCATTTGTTTGTAGTAAATCGGAGGTGAAGATGGAGATATTCATCGAGGTTTTCAGTCA 275  
 Db 240 ATATCTTTCTTCAATAAAGTGGCGGT---GGTGGAGACATCCACAGGGCTTCCAGTCT 296  
 QY 276 CTTCTGTTGCAATTAACAGAACTGACACTGAATATGTGCTTAGAAGTGCACACGGGCTC 335  
 Db 297 CTTCTACCCGAAGTGAACAGACTGGCAGCAGTACTTGTCTAGGGTGGCCACAGGCTC 356  
 QY 336 TTTGGAGAAAAGTCTTATGATTTTCCACAGGTTTACAGATTTCTGTGSCAAATTTCTAC 395  
 Db 357 TTTGGGGAAGTCTTGTGATTTTCTCTCATCTTTTAGAGATTTCTGCGCAAAATTTCTAC 416  
 QY 396 CAAGCAACGATAAAACAGCTAGACTTTGTGAATGATACAGAGAGTCCACAAACGCTGTA 455  
 Db 417 CAAGCAGATGGAGAGCTTGACTTTATCAGCGCGTAGAGAGTCCAGAAAACACATA 476  
 QY 456 AACTCTCTGGGTGCTGATAAAACTAAAG 483  
 Db 477 AACACCTGGGTAGCTGAAAAGACAGAG 504  
 RESULT 7  
 AX147509 1324 bp DNA linear PAT 08-JUN-2001  
 LOCUS Sequence 63 from Patent WO0136632.  
 DEFINITION AX147509  
 ACCESSION AX147509  
 VERSION AX147509.1 GI:14346666  
 KEYWORDS human.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1324)  
 AUTHORS Levine, Z., David, A., Azar, I., Khosravi, R. and Bernstein, J.  
 TITLE Variants of alternative splicing  
 JOURNAL Patent: WO 0136632-A 63 25-MAY-2001;  
 CompuGen Ltd. (IL)  
 FEATURES  
 Location/Qualifiers  
 source  
 1..1324  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 364 a 287 c 360 g 313 t  
 ORIGIN  
 Query Match 35.0%; Score 246.8; DB 6; Length 1324;  
 Best Local Similarity 75.8%; Pred. No. 9.1e-60;  
 Matches 332; Conservative 0; Mismatches 102; Indels 4; Gaps 2;  
 QY 47 CCGTCATGATGCTCTATCAGAACCAATGCCACATTTGCATTAACCTTTTGAAGAACG 106  
 Db 204 CCATCATGATGTTCTCCAGAGCAAAATGGACCTTTTGCCCTTAACCTTTTGAAGAACG 263  
 QY 107 TAGGGGAAAAACAACCTCAAAACAACCTTATTTTTTTC-CCCATGAGCATATCATCAGCCTTG 165  
 Db 264 TGGGTAAGACAACTCGAAGATGTGTTTTTCTCACCCTAGCATGCTGTGTCCTGG 323  
 QY 166 CCATGGTTTTCATGGGGCAAGGAAACACTGAGCTCAGATGTCTCAGGACACTTTGT 225  
 Db 324 CCATGGTCTACATGGGGCAAGGAAACACCGCTGCACAGATGGCCCAAGTACTTCT 383  
 QY 226 TTAGTAAATCGAGGTGAAGATGAGATATTCATCGAGGTTTTCAGTCACTTCTTTGTTG 285  
 Db 384 TCAATAAAGTGGCGGT---GGTGGAGACATCCACAGGGCTTCCAGTCTCTTCTCACCG 440  
 QY 286 CAATTAACAGAACTGACACTGAATATGTCTTAGAACTGCCAACGGGCTCTTTGGAGAA 345

Db 441 AAGTGAACAGACTGGCAGCAGTACTTGGTGGCCCAACAGGCTCTTTGGGAAA 500  
Qy 346 AGTCTTATGATTTCTCAGAGTTTACAGATTCCTGTGGCAAAATTTACCAAGCAAG 405  
Db 501 AGTCTTGTGATTTCTCTCATCTTTTAGAGATTCCTGCCAAAATTTCTACCAAGCAGA 560  
Qy 406 TAAACAGCTAGACTTTGTGAATGATACAGAGAGTCCACAAACAGTGTAACTCCCTGGG 465  
Db 561 TGGAGGAGCTTGACTTTATCAGCGCGTAGAGAGTCCAGAAAACATAAACACCTGGG 620  
Qy 466 TTGCTGATAAACTAAAG 483  
Db 621 TAGCTGAACAGACAGAAG 638

RESULT 8  
AX147468  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

AX147468 1349 bp DNA linear PAT 08-JUN-2001  
Sequence 22 from Patent WO0136632.  
AX147468  
Version AX147468.1 GI:14346625  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1349)  
Levine, Z., David, A., Azar, I., Khosravi, R. and Bernstein, J.  
Variants of alternative splicing  
Patent: WO 0136632-A 22 25-MAY-2001;  
Compugen Ltd. (IL)  
Location/Qualifiers  
1. 1349  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
370 a 358 g 323 t

Query Match 35.0%; Score 246.8; DB 6; Length 1349;  
Best Local Similarity 75.8%; Pred. No. 9.1e-60;  
Matches 332; Conservative 0; Mismatches 102; Indels 4; Gaps 2;  
Qy 47 CCGTCATGATGCTCTATCAGAGCAAAATGGCACAATTCGATTAACCTTTGAAAAGC 106  
Db 204 CCATCATGATGCTCTCGCAGAGCAAAATGGCACCCTTTGGCCTTAAACCTTTGAAAAGC 263  
Db 107 TAGGGGAAAACAACTCAAACTTATTTTTC-CCCATGAGCATATCATCAGCCTTGG 165  
Db 264 TGGGTAAGACAACCTCGAAGATGTGTTTCTCACCATGAGCATGCTGTGCCCTGG 323  
Qy 166 CCATGTTTTCATGGGGCAAGGAAACACTGCAGCTCAGATGCTCAGGCATTTGTT 225  
Db 324 CCATGTTCTACATGGGGCAAGGAAACACCGCTGCACAGATGGCCAGATACCTTTCT 383  
Qy 226 TTAGTAAATCGAGGTGAAGATGAGATATTCATCGAGGTTTTCAGTCACCTTTCTT 285  
Db 384 TCAATAAAGTGGCGGT---GGTGAGACATCCACAGGCTTCCAGTCTCTTCTCACCG 440  
Qy 286 CAATTAACAGACTGACACTGAATATGTGTTAGAACTGCCAAGGGCTCTTTGGAGAA 345  
Db 441 AAGTGAACAGACTGGCAGCAGTACTTGTCTAGGTTGGCCCAACAGGCTCTTTGGGAAA 500  
Qy 346 AGTCTTATGATTTCTCAGAGTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCAAG 405  
Db 501 AGTCTTGTGATTTCTCTCATCTTTTAGAGATTCCTGCCAAAATTTCTACCAAGCAGA 560  
Qy 406 TAAACAGCTAGACTTTGTGAATGATACAGAGAGTCCACAAACAGTGTAACTCCCTGGG 465  
Db 561 TGGAGGAGCTTGACTTTATCAGCGCGTAGAGAGTCCAGAAAACATAAACACCTGGG 620  
Qy 466 TTGCTGATAAACTAAAG 483

RESULT 10  
AX140381  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

AX140381 1465 bp DNA linear PAT 14-JUN-2002  
Sequence 3028 from Patent WO0229103.  
AX140381  
Version AX140381.1 GI:21443086  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1410)  
Levine, Z., David, A., Azar, I., Khosravi, R. and Bernstein, J.  
Variants of alternative splicing  
Patent: WO 0136632-A 62 25-MAY-2001;  
Compugen Ltd. (IL)  
Location/Qualifiers  
1. 1410  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
376 a 310 c 390 g 333 t 1 others

Query Match 35.0%; Score 246.8; DB 6; Length 1410;  
Best Local Similarity 75.8%; Pred. No. 9.1e-60;  
Matches 332; Conservative 0; Mismatches 102; Indels 4; Gaps 2;  
Qy 47 CCGTCATGATGCTCTATCAGAGCAAAATGGCACAATTCGATTAACCTTTGAAAAGC 106  
Db 204 CCATCATGATGCTCTCGCAGAGCAAAATGGCACCCTTTGGCCTTAAACCTTTGAAAAGC 263  
Db 107 TAGGGGAAAACAACTCAAACTTATTTTTC-CCCATGAGCATATCATCAGCCTTGG 165  
Db 264 TGGGTAAGACAACCTCGAAGATGTGTTTCTCACCATGAGCATGCTGTGCCCTGG 323  
Qy 166 CCATGTTTTCATGGGGCAAGGAAACACTGCAGCTCAGATGCTCAGGCATTTGTT 225  
Db 324 CCATGTTCTACATGGGGCAAGGAAACACCGCTGCACAGATGGCCAGATACCTTTCT 383  
Qy 226 TTAGTAAATCGAGGTGAAGATGAGATATTCATCGAGGTTTTCAGTCACCTTTCTT 285  
Db 384 TCAATAAAGTGGCGGT---GGTGAGACATCCACAGGCTTCCAGTCTCTTCTCACCG 440  
Qy 286 CAATTAACAGACTGACACTGAATATGTGTTAGAACTGCCAAGGGCTCTTTGGAGAA 345  
Db 441 AAGTGAACAGACTGGCAGCAGTACTTGTCTAGGTTGGCCCAACAGGCTCTTTGGGAAA 500  
Qy 346 AGTCTTATGATTTCTCAGAGTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCAAG 405  
Db 501 AGTCTTGTGATTTCTCTCATCTTTTAGAGATTCCTGCCAAAATTTCTACCAAGCAGA 560  
Qy 406 TAAACAGCTAGACTTTGTGAATGATACAGAGAGTCCACAAACAGTGTAACTCCCTGGG 465  
Db 561 TGGAGGAGCTTGACTTTATCAGCGCGTAGAGAGTCCAGAAAACATAAACACCTGGG 620  
Qy 466 TTGCTGATAAACTAAAG 483

RESULT 10  
AX140381  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

AX140381 1465 bp DNA linear PAT 14-JUN-2002  
Sequence 3028 from Patent WO0229103.  
AX140381  
Version AX140381.1 GI:21443086  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1410)  
Levine, Z., David, A., Azar, I., Khosravi, R. and Bernstein, J.  
Variants of alternative splicing  
Patent: WO 0136632-A 62 25-MAY-2001;  
Compugen Ltd. (IL)  
Location/Qualifiers  
1. 1410  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
376 a 310 c 390 g 333 t 1 others

Db 621 TAGCTGAACAGACAGAAG 638  
RESULT 9  
AX147508  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

AX147508 1410 bp DNA linear PAT 08-JUN-2001  
Sequence 62 from Patent WO0136632.  
AX147508  
Version AX147508.1 GI:14346665  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1410)  
Levine, Z., David, A., Azar, I., Khosravi, R. and Bernstein, J.  
Variants of alternative splicing  
Patent: WO 0136632-A 62 25-MAY-2001;  
Compugen Ltd. (IL)  
Location/Qualifiers  
1. 1410  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
376 a 310 c 390 g 333 t 1 others

Query Match 35.0%; Score 246.8; DB 6; Length 1410;  
Best Local Similarity 75.8%; Pred. No. 9.1e-60;  
Matches 332; Conservative 0; Mismatches 102; Indels 4; Gaps 2;  
Qy 47 CCGTCATGATGCTCTATCAGAGCAAAATGGCACAATTCGATTAACCTTTGAAAAGC 106  
Db 204 CCATCATGATGCTCTCGCAGAGCAAAATGGCACCCTTTGGCCTTAAACCTTTGAAAAGC 263  
Db 107 TAGGGGAAAACAACTCAAACTTATTTTTC-CCCATGAGCATATCATCAGCCTTGG 165  
Db 264 TGGGTAAGACAACCTCGAAGATGTGTTTCTCACCATGAGCATGCTGTGCCCTGG 323  
Qy 166 CCATGTTTTCATGGGGCAAGGAAACACTGCAGCTCAGATGCTCAGGCATTTGTT 225  
Db 324 CCATGTTCTACATGGGGCAAGGAAACACCGCTGCACAGATGGCCAGATACCTTTCT 383  
Qy 226 TTAGTAAATCGAGGTGAAGATGAGATATTCATCGAGGTTTTCAGTCACCTTTCTT 285  
Db 384 TCAATAAAGTGGCGGT---GGTGAGACATCCACAGGCTTCCAGTCTCTTCTCACCG 440  
Qy 286 CAATTAACAGACTGACACTGAATATGTGTTAGAACTGCCAAGGGCTCTTTGGAGAA 345  
Db 441 AAGTGAACAGACTGGCAGCAGTACTTGTCTAGGTTGGCCCAACAGGCTCTTTGGGAAA 500  
Qy 346 AGTCTTATGATTTCTCAGAGTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCAAG 405  
Db 501 AGTCTTGTGATTTCTCTCATCTTTTAGAGATTCCTGCCAAAATTTCTACCAAGCAGA 560  
Qy 406 TAAACAGCTAGACTTTGTGAATGATACAGAGAGTCCACAAACAGTGTAACTCCCTGGG 465  
Db 561 TGGAGGAGCTTGACTTTATCAGCGCGTAGAGAGTCCAGAAAACATAAACACCTGGG 620  
Qy 466 TTGCTGATAAACTAAAG 483  
Db 621 TAGCTGAACAGACAGAAG 638

RESULT 10  
AX140381  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

AX140381 1465 bp DNA linear PAT 14-JUN-2002  
Sequence 3028 from Patent WO0229103.  
AX140381  
Version AX140381.1 GI:21443086  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1410)  
Levine, Z., David, A., Azar, I., Khosravi, R. and Bernstein, J.  
Variants of alternative splicing  
Patent: WO 0136632-A 62 25-MAY-2001;  
Compugen Ltd. (IL)  
Location/Qualifiers  
1. 1410  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
376 a 310 c 390 g 333 t 1 others

Query Match 35.0%; Score 246.8; DB 6; Length 1410;  
Best Local Similarity 75.8%; Pred. No. 9.1e-60;  
Matches 332; Conservative 0; Mismatches 102; Indels 4; Gaps 2;  
Qy 47 CCGTCATGATGCTCTATCAGAGCAAAATGGCACAATTCGATTAACCTTTGAAAAGC 106  
Db 204 CCATCATGATGCTCTCGCAGAGCAAAATGGCACCCTTTGGCCTTAAACCTTTGAAAAGC 263  
Db 107 TAGGGGAAAACAACTCAAACTTATTTTTC-CCCATGAGCATATCATCAGCCTTGG 165  
Db 264 TGGGTAAGACAACCTCGAAGATGTGTTTCTCACCATGAGCATGCTGTGCCCTGG 323  
Qy 166 CCATGTTTTCATGGGGCAAGGAAACACTGCAGCTCAGATGCTCAGGCATTTGTT 225  
Db 324 CCATGTTCTACATGGGGCAAGGAAACACCGCTGCACAGATGGCCAGATACCTTTCT 383  
Qy 226 TTAGTAAATCGAGGTGAAGATGAGATATTCATCGAGGTTTTCAGTCACCTTTCTT 285  
Db 384 TCAATAAAGTGGCGGT---GGTGAGACATCCACAGGCTTCCAGTCTCTTCTCACCG 440  
Qy 286 CAATTAACAGACTGACACTGAATATGTGTTAGAACTGCCAAGGGCTCTTTGGAGAA 345  
Db 441 AAGTGAACAGACTGGCAGCAGTACTTGTCTAGGTTGGCCCAACAGGCTCTTTGGGAAA 500  
Qy 346 AGTCTTATGATTTCTCAGAGTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCAAG 405  
Db 501 AGTCTTGTGATTTCTCTCATCTTTTAGAGATTCCTGCCAAAATTTCTACCAAGCAGA 560  
Qy 406 TAAACAGCTAGACTTTGTGAATGATACAGAGAGTCCACAAACAGTGTAACTCCCTGGG 465  
Db 561 TGGAGGAGCTTGACTTTATCAGCGCGTAGAGAGTCCAGAAAACATAAACACCTGGG 620  
Qy 466 TTGCTGATAAACTAAAG 483  
Db 621 TAGCTGAACAGACAGAAG 638

RESULT 10  
AX140381  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

AX140381 1465 bp DNA linear PAT 14-JUN-2002  
Sequence 3028 from Patent WO0229103.  
AX140381  
Version AX140381.1 GI:21443086  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1410)  
Levine, Z., David, A., Azar, I., Khosravi, R. and Bernstein, J.  
Variants of alternative splicing  
Patent: WO 0136632-A 62 25-MAY-2001;  
Compugen Ltd. (IL)  
Location/Qualifiers  
1. 1410  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
376 a 310 c 390 g 333 t 1 others

Query Match 35.0%; Score 246.8; DB 6; Length 1410;  
Best Local Similarity 75.8%; Pred. No. 9.1e-60;  
Matches 332; Conservative 0; Mismatches 102; Indels 4; Gaps 2;  
Qy 47 CCGTCATGATGCTCTATCAGAGCAAAATGGCACAATTCGATTAACCTTTGAAAAGC 106  
Db 204 CCATCATGATGCTCTCGCAGAGCAAAATGGCACCCTTTGGCCTTAAACCTTTGAAAAGC 263  
Db 107 TAGGGGAAAACAACTCAAACTTATTTTTC-CCCATGAGCATATCATCAGCCTTGG 165  
Db 264 TGGGTAAGACAACCTCGAAGATGTGTTTCTCACCATGAGCATGCTGTGCCCTGG 323  
Qy 166 CCATGTTTTCATGGGGCAAGGAAACACTGCAGCTCAGATGCTCAGGCATTTGTT 225  
Db 324 CCATGTTCTACATGGGGCAAGGAAACACCGCTGCACAGATGGCCAGATACCTTTCT 383  
Qy 226 TTAGTAAATCGAGGTGAAGATGAGATATTCATCGAGGTTTTCAGTCACCTTTCTT 285  
Db 384 TCAATAAAGTGGCGGT---GGTGAGACATCCACAGGCTTCCAGTCTCTTCTCACCG 440  
Qy 286 CAATTAACAGACTGACACTGAATATGTGTTAGAACTGCCAAGGGCTCTTTGGAGAA 345  
Db 441 AAGTGAACAGACTGGCAGCAGTACTTGTCTAGGTTGGCCCAACAGGCTCTTTGGGAAA 500  
Qy 346 AGTCTTATGATTTCTCAGAGTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCAAG 405  
Db 501 AGTCTTGTGATTTCTCTCATCTTTTAGAGATTCCTGCCAAAATTTCTACCAAGCAGA 560  
Qy 406 TAAACAGCTAGACTTTGTGAATGATACAGAGAGTCCACAAACAGTGTAACTCCCTGGG 465  
Db 561 TGGAGGAGCTTGACTTTATCAGCGCGTAGAGAGTCCAGAAAACATAAACACCTGGG 620  
Qy 466 TTGCTGATAAACTAAAG 483



ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1523)
TITLE	Strausberg, R.
JOURNAL	Direct Submission
REMARK	Submitted (12-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
COMMENT	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Sequencing Center (NISC), National Institutes of Health Intramural Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a> Contact: <a href="mailto:nisc.mgc@nhgri.nih.gov">nisc.mgc@nhgri.nih.gov</a> Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tongson, E.E., Touchman, J.W., Taurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.
FEATURES	Location/Qualifiers 1..1523 /organism="Homo sapiens" /db_xref="LOCUSID:5269" /db_xref="taxon:9606" /clone="MGC:2180 IMAGE:3051381" /tissue_type="Colon, adenocarcinoma" /clone_lib="NIH MGC 15" /lab_host="DH10B-R" /note="Vector: pOTB7" 235..1365 /codon_start=1 /product="serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 6" /protein_id="AAH01394.1" /db_xref="GI:12655087" /translation="MDVLAERNGTFALNLLKTLGKDNKNVFFSPMSMCSALAMVYMG AKGNTAAQALSLFNKSGGGDIHQGFOSLLTEVNTKTQYLLRVANRLFEKSCDF LSPFSDCKQFYQAESELDLFI SAVERKHINTVWAEKTEKIAELLSPGSDPLTR LVLNVAVPRGNWDQFQKENTFEERLFKVSKEEKFQVMPKQSTFKYTGIEIPTOI LVLVYVGLNELMIIMLPDITDLRTVEKELTYEKFVWTRLDMDDEEREVEVSLPRFKL EESYDMSVRLNGLMTDFELGKADFGSGMSOTDLSLSKVHKSFPVNEEGTEAAAT AAIMWMGRFVPRCADHPFLFFQHKSTGILFLCGRFSP"
CDS	407 a 351 c 405 g 360 t
BASE COUNT	407 a 351 c 405 g 360 t
ORIGIN	
Query Match	35.0%; Score 246.8; DB 9; Length 1523;
Best Local Similarity	75.8%; Pred. No. 9.1e-60;
Matches 332; Conservative	0; Mismatches 102; Indels 4; Gaps 2;
QY	47 CGTCATCGATGCTCTATCAGAGCAATGGCATTGTCATTAAACCTTTGAAAAGC 106
Db	230 CCATCGATGATGTTCTCCAGAGCAAAATGGCACCTTGGCTTAAACCTTTGAAAAGC 289
QY	107 TAGGGGAAAAACAACCTCAACAACTTATTTTTC-CCCATGAGCATATCATCGCCTTGG 165
Db	290 TGGGTAAGACAACTCGAAGATGTGTTTTCTACCCATGAGCATGTCCTGTCCTGG 349
QY	166 CCATGGTTTTTCATGGGGGCAAGGAAACACTGCAGCTCAGATGCTCTCAGCAGCACTTTGTT 225
Db	350 CCATGGTTTACATGGGGGCAAGGAAACACCGCTGCACAGATGCCACGATACCTTCTT 409
QY	226 TTAGTAAATCGAGGTGAAGATCGAGATATTCATCGAGGTTTTTCAGTCACTTCTTTGTTG 285
Db	410 TCAATAAAGAGTGGCGT---GGTGGAGACATCCACAGGGCTTCAGATCTCTTCTACCG 466
QY	286 CAATTAAACGAACTGACACTGAATATGTGTTAGAACTGCCAACGGGCTCTTTGGAGAAA 345
Db	467 AAGTGAACAAAGACTGGCAGCAGTACTTGCCTTAGGGTGGCCAAACAGGCTCTTTGGGAAA 526
QY	346 AGCTTTATGATTCTCTCAGAGTTTTCAGAGATTTCTGTGGCAATTTTACCAGCAAGA 405
Db	527 AGCTTTGATTTCTCTCATCTTTAGAGATTTCTGCCAAAATTTTACCAGCAGAGA 586
QY	406 TAAACACAGCTAGACTTTTGTGAATATACAGAGAAAGTCCACACACGTTAAACTCCTGGG 465
Db	587 TGGAGGAGCTTGACTTTATCAGCGCGTAGAGAGTCCAGAAACACATATAACACCTGGG 646
QY	466 TTGCTGATAAACTAAAG 483
Db	647 TAGCTGAAAAAGACAGAAG 664
RESULT 13	
AK094983	
LOCUS	AK094983.1 GI:21754158
DEFINITION	Homo sapiens CDNA FLJ37664 fis, clone BRHIP2011491, highly similar to PLACENTAL THROMBIN INHIBITOR.
ACCESSION	AK094983
VERSION	AK094983.1
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens hippocampus CDNA to mRNA, clone BRHIP2011491.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2036)
AUTHORS	Isogai, T. and Yamamoto, J.
TITLE	Direct Submission
JOURNAL	Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazuo-Kamata, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; CDNA full insert sequencing: Research Association for Biotechnology (RAB); CDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES	Location/Qualifiers 1..2036 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="BRHIP2011491" /tissue_type="hippocampus" /clone_lib="BRHIP2" /note="cloning vector: pME185FLJ"
BASE COUNT	459 a 564 c 535 g 478 t
ORIGIN	

Evaluation; clone selection for full insert sequencing: RAB and

HRI. Location/Qualifiers  
1. 3271  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="SPLEN2000247"  
/tissue type="spleen"  
/clone\_lib="SPLEN2"  
/notes="cloning vector: pME18SFL3"  
BASE COUNT 796 a 828 c 799 g 848 t  
ORIGIN  
Query Match 35.0%; Score 246.8; DB 9; Length 3271;  
Best Local Similarity 75.8%; Pred. No. 9.2e-60;  
Matches 332; Conservative 0; Mismatches 102; Indels 4; Gaps 2;  
QY 47 CCGTCATGATGCTCTATCAGAAAGCAATGGCAGATTTGGCATTAACCTTTTGAAGAAGC 106  
DB 774 CCATCATGATGTTCTCGCAGAAAGCAATGGCAGATTTGGCATTAACCTTTTGAAGAAGC 833  
QY 107 TAGGGGAAACCACTCAAACTTATTTTTC-CCCATGAGCATATCATCAGCCTTGG 165  
DB 834 TGGGTAAAGCAACTCGAAGATGTTTCTCACCATGAGCATGCTGTGCTGCTGG 893  
QY 166 CCATGTTTTTCATGGGGGCAAGGAAACACATGAGCTCAGATGCTCAGGACCTTTGT 225  
DB 894 CCATGTTCTACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 953  
QY 226 TTAGTAAATCGAGGTGAAGATGAGATATTCATCGAGTTTTCAGTCACCTTTCTGTTG 285  
DB 954 TCAATAAAGTGGCGT---GGTGGAGACATCCACAGGGCTTCAGTCTCTTCTCACCG 1010  
QY 286 CAATTAAACGAACCTGACACTGAATATGCTTTAGAACTGCCAACGGGCTTTTGGAGAAA 345  
DB 1011 AAGTGAACAGACTGCGACGCTGCTTCTTAGGATGGCCAAACAGGCTCTTTTGGGAAA 1070  
QY 346 AGTCTTATGATTTCTCAGAGTTTACAGATTCCTGTGGCAATTCCTACCAAGCAACGA 405  
DB 1071 AGTCTTGTGATTTCTCTCATCTTTTAGAGATTCCTGCCAAATTTCTACCAAGCAGAG 1130  
QY 406 TAAACAGCTAGACTTTGTGAATGATACAGAGAAAGTCCCAACACGCTGTAACCTCTGGG 465  
DB 1131 TGGAGAGCTTGACTTTTACAGCGCGTAGAGAGTCCAGAAACACATAAACACCTGGG 1190  
QY 466 TTGCTGATAAACTAAAG 483  
DB 1191 TAGCTGAAAAGACAGAAG 1208

RESULT 15  
AK057138  
LOCUS  
DEFINITION  
Homo sapiens cDNA FLJ32576 fis, clone SPLEN2000247, highly similar  
to PLACENTAL THROMBIN INHIBITOR.  
ACCESSION  
AK057138.1 GI:16552727.  
VERSION  
oligo capping; fis (full insert sequence).  
KEYWORDS  
Homo sapiens spleen cDNA to mRNA, clone\_lib:SPLEN2  
clone:SPLEN2000247.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1  
AUTHORS  
Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,  
Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J.,  
Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,  
Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,  
Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,  
Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahara,K.,  
Maeuho,Y., Nagai,K. and Isogai,T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 3271)  
Isogai,T., Otsuki,T. and Sugiyama,T.  
Direct Submission  
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and

RESULT 15  
AK027149  
LOCUS  
DEFINITION  
Homo sapiens cDNA: FLJ23496 fis, clone LNG02535, highly similar to  
S69272 human mRNA for cytoplasmic antiproteinase-38 kDa  
intracellular serine proteinase inhibitor.  
ACCESSION  
AK027149  
VERSION  
oligo capping; fis (full insert sequence).  
KEYWORDS  
Homo sapiens human lung cDNA to mRNA, clone\_lib:LNG clone:LNG02535.  
SOURCE  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (slices)  
AUTHORS  
Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,  
Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,  
Fujiiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hiro,M.,  
Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,  
Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2759)  
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,



**TITLE**  
JOURNAL

**COMMENT**

**FEATURES**  
source

**misc\_feature**

**BASE COUNT**  
ORIGIN

**Location/Qualifiers**  
1..2759  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="LNG02535"  
/tissue\_type="human lung"  
/clone\_lib="LNG"  
/note="cloning vector pME18SFL3"  
1..2759  
/note="highly similar to S69272 human mRNA for cytoplasmic  
antiproteinase-38 kda intracellular serine proteinase  
inhibitor"

Shibahara, T., Tanaka, T. and Nakamura, Y.  
Direct Submission  
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,  
University of Tokyo, Laboratory of Genome Structure Analysis, Human  
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,  
Japan (E-mail: cdnal@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,  
Fax: 81-3-5449-5416)  
NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert  
sequencing: Research Association for Biotechnology; cDNA library  
construction, 5'- & 3'-end one pass sequencing: Department of  
Virology and Human Genome Center, Institute of Medical Science,  
University of Tokyo (partly supported by Science and Technology  
Agency).

Query Match 34.7%; Score 245.2; DB 9; Length 2759;  
Best Local Similarity 75.6%; Pred. No. 2.7e-55;  
Matches 331; Conservative 0; Mismatches 103; Indels 4; Gaps 2;

QY 47 CCGTCATGGATGCTCTATCAGAACCAATGGCACATTTGCATTAAACCTTTTGAAAAAGC 106  
DB 1465 CCATCATGGATGTTCTCGCAGAGCAATGGCACCTTTGGCTTAACCTTTTGAAAAACAC 1524

QY 107 TAGGGGAAACAACTCAAAACCTTATTTTTC-CCCATGAGCATATCATCAGCCTTGG 165  
DB 1525 TGGGTAAAGACAACTCGAGAGATGTGTTTCTCACCCTGAGCATGCTGTGCCCTGG 1584

QY 166 CCATGGTTTTCATGGGGGCAAGGAAACACTGCAGCTCAGATGTCTCAGGCACCTTTGTT 225  
DB 1585 CCATGGTCTACATGGGGGCAAGGAAACACCCGCTGCACAGATGGCCAGATACCTTCTT 1644

QY 226 TTAGTAAATCGGAGGTGAGATGGAGATATTCATCGAGGTTTTCAGTCACCTTCTGTTG 285  
DB 1645 TCAATAAAGTGGCGT---GGTGGAGACATCCACAGGGCTTCCAGTCTCTTCTCACCG 1701

QY 286 CAATTAACAGAACTGACACTGAATATGTCTTAGAACTGCCAACGGCTCTTTGGAGAA 345  
DB 1702 AAGTGAACAAAGCTGGCAGCAGTACTTGTAGGATGCCAACAGGCTCTTTGGGGAA 1761

QY 346 AGCTTTATGATTTTCCCTCAGAGTTTACAGATTCCTGTGGCAAAATTTACCAAGCAAG 405  
DB 1762 AGCTTGTGATTTTCTCTCATCTTTTAGAGATTTCTGCCAAAAATTTACCAAGCGGAG 1821

QY 406 TAAACAGCTAGACTTTTGTGAATGATACAGAGAAGTCCACACAGCTGTAAACTCTGGG 465  
DB 1822 TGGAGGAGCTTGACTTTTATCAGCGCGGTAGAGAGTCCAGAAAAACACATAAACACCTGGG 1881

QY 466 TTCTGTATAAACTAAAG 483  
DB 1882 TAGCTGAAAAGACAGAG 1899

Search completed: May 21, 2003, 17:33:27  
Job time : 2178 secs



Result No.	Query	Score	Query			Description
			Match	Length	ID	
1	704	99.7	706	21	Nucleotide	
2	704	99.7	706	23	Human serif	
3	586	83.0	599	21	Nucleotide	
4	586	83.0	599	23	Human serif	
5	577	81.7	608	21	Nucleotide	
6	577	81.7	608	23	Human serif	
7	462.8	65.6	943	24	Human novel	
8	251.2	35.6	1624	21	Human protein	
9	230.4	35.5	695	22	Angiotensin	

PPF 29-FEB-2000; 2000WO-US05082.

XX 01-MAR-1999; 99US-0122276.  
PR 12-MAR-1999; 99US-0124094.  
PR 18-AUG-1999; 99US-0149452.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Ruben SM, Ni J;  
PI WPI; 2000-579284/54.  
DR P-PSDB; AAB08776.  
XX Novel human serine protease inhibitor superfamily of proteins useful  
PT for diagnosing, treating, preventing and/or prognosing disorders  
PT related to the serpin polypeptides  
XX Claim 1; Page 202; 215pp; English.  
XX The present sequence encodes a human serpin polypeptide. Serpin is  
CC a member of the serine protease inhibitor superfamily of polypeptides.  
CC The serpin polynucleotides and polypeptides are used to prevent, treat  
CC or ameliorate a medical condition in e.g. humans, mice, rabbits, goats,  
CC horses, cats, dogs, chickens or sheep. Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiogenesis, nervous system disorders,  
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
CC fungi and ocular disorders e.g. corneal infection. The polypeptides can  
CC also be used to aid wound healing and epithelial cell proliferation,  
CC to prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities.  
XX Query Match 99.78; Score 704; DB 21; Length 706;  
XX Best Local Similarity 100.0%; Pred. NO. 8.3e-199; Indels 0; Gaps 0;  
XX Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGCACGAGCTTCGCTCTGGGACGCTGGGGGAGAACTGGGGCTCACCGTCATGGATGCT 60  
DB 1 GGCACGAGCTTCGCTCTGGGACGCTGGGGGAGAACTGGGGCTCACCGTCATGGATGCT 60  
QY 61 CTATCAGAAGCAATGGCAGATTTGCAATTAACCTTTTGAAGAGCTAGGGAAACAAC 120  
DB 61 CTATCAGAAGCAATGGCAGATTTGCAATTAACCTTTTGAAGAGCTAGGGAAACAAC 120  
QY 121 TCAACAACTTATTTTTCCTCCCATGAGCATATCATCAGCCTTGGCCATGTTTTCATGG 180  
DB 121 TCAACAACTTATTTTTCCTCCCATGAGCATATCATCAGCCTTGGCCATGTTTTCATGG 180  
QY 181 GGGCAAGGGAACACATCGAGTCAGATGTCAGGCACTTTGTTAGTAAATCGGAG 240  
DB 181 GGGCAAGGGAACACATCGAGTCAGATGTCAGGCACTTTGTTAGTAAATCGGAG 240  
QY 241 GTGAAGATGGAGATTTTCATCGAGTTTTCAGTCACCTTTGTTGCAATTAACAGAACTG 300  
DB 241 GTGAAGATGGAGATTTTCATCGAGTTTTCAGTCACCTTTGTTGCAATTAACAGAACTG 300  
QY 301 ACACCTGAATATGCTGTAGAACTGCCAACGGGCTTTTGGAGAAAGCTTTATGATTTC 360  
DB 301 ACACCTGAATATGCTGTAGAACTGCCAACGGGCTTTTGGAGAAAGCTTTATGATTTC 360  
QY 361 TCACAGTTTTHACAGATTCCTGTGCAATTTCTACCAAGCAACGATTAACAGCTAGACT 420  
DB 361 TCACAGTTTTHACAGATTCCTGTGCAATTTCTACCAAGCAACGATTAACAGCTAGACT 420  
QY 421 TTGTGAATCATACAGAGAACTGCCAACACGCTGTAAACTCTCTGGGTTGCTGATAAACTA 480  
DB 421 TTGTGAATCATACAGAGAACTGCCAACACGCTGTAAACTCTCTGGGTTGCTGATAAACTA 480

QY 481 AAGCCTGGAAATTTATTCAACAAGCCTCTCACATCTGAGGAGCCAGGAATCGCCTCTT 540  
DB 481 AAGCCTGGAAATTTATTCAACAAGCCTCTCACATCTGAGGAGCCAGGAATCGCCTCTT 540  
QY 541 CCTCTTTGTTACTGCAAAAGCCTGCTTTTCACAGCCCTACTGTTTCTATTTCCCAAT 600  
DB 541 CCTCTTTGTTACTGCAAAAGCCTGCTTTTCACAGCCCTACTGTTTCTATTTCCCAAT 600  
QY 601 GCAACTCTCTGACCCCGCATGGCATGTGGTGNCTCCCTCCCTGTGAGCAGNTGTGA 660  
DB 601 GCAACTCTCTGACCCCGCATGGCATGTGGTGNCTCCCTCCCTGTGAGCAGNTGTGA 660  
QY 661 CTATTAACCTGCCGCAATTTTCATCTGTAAATAAAAAA 706  
DB 661 CTATTAACCTGCCGCAATTTTCATCTGTAAATAAAAAA 706  
RESULT 2  
AAL42746  
ID AAL42746 standard; cDNA; 706 BP.  
XX AAL42746;  
AC AAL42746;  
XX 19-JUL-2002 (first entry)  
DT Human serine protease inhibitor (serpin) coding sequence HMCIS41.  
XX Human; serine protease inhibitors; serpin; blood coagulation;  
KW proteinase-activated physiological process; gene therapy;  
KW antitense therapy; cancer; inflammation; neurological disease;  
KW immune disease; gene; ss; HMCIS41.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH 3,650  
FT /\*tag= a  
FT /product= "Human serpin protein HMCIS41"  
FT /transl\_except= (pos:632..635, aa:Xaa)  
FT /note= "No start codon is given; Xaa is unknown"  
XX -WO200155390-A1.  
PN 02-AUG-2001.  
XX 26-JAN-2001; 2001WO-US02484.  
XX 28-JAN-2000; 2000US-0178769.  
PR 29-FEB-2000; 2000WO-US05082.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Ni J, Ruben SM, Shi Y;  
XX WPI; 2001-457730/49.  
DR P-PSDB; AAO14914.  
XX Isolated serine protease inhibitor polypeptide for treating, preventing  
PT and/or prognosing cancers, inflammation, immune disorders and  
PT neurological diseases and also for testing and detection e.g. diagnosis  
PT  
XX Claim 1; Page 273; 288pp; English.  
XX The invention comprises the amino acid and coding sequences of a group of  
CC novel human serine protease inhibitors (serpins). Serpins are a  
CC superfamily of single chain proteins that contain a conserved structure  
CC of approximately 370-420 amino acids and generally range between 50 and  
CC 100 kDa in molecular mass. The majority of serpins function as protease  
CC inhibitors and so are involved in the regulation of several proteinase-  
CC activated physiological processes (e.g. blood coagulation). The Serpin

CC DNA and amino acid sequences of the invention are useful in the treatment  
 CC (e.g. gene therapy, antisense therapy) and diagnosis of conditions such  
 CC as cancer, inflammation, neurological diseases and immune diseases. The  
 CC present cDNA sequence represents a human serpin coding sequence of the  
 CC invention.  
 XX  
 SQ Sequence 706 BP; 207 A; 166 C; 148 G; 183 T; 2 other;

Query Match 99.7%; Score 704; DB 23; Length 706;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-199;  
 Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGCACGAGCTTCCTCTCTGGGAGCTCGCGGAGAACTGGGGCTACCGCTATGATGCT 60  
 DB 1 GGCACGAGCTTCCTCTCTGGGAGCTCGCGGAGAACTGGGGCTACCGCTATGATGCT 60  
 QY 61 CTATCAGAACAAATGGCAGATTTGCATTAAACCTTTTGAAGAAGCTAGGGAAACAC 120  
 DB 61 CTATCAGAACAAATGGCAGATTTGCATTAAACCTTTTGAAGAAGCTAGGGAAACAC 120  
 QY 121 TCAACAACTTATTTTCCCATGAGCATATCATCAGCCTTGGCCATGGTTTTCATGG 180  
 DB 121 TCAACAACTTATTTTCCCATGAGCATATCATCAGCCTTGGCCATGGTTTTCATGG 180  
 QY 181 GGGCAAGGAGAACTCGAGCTCAGATGCTCAGCAGCACTTTGTTAGTAAATCGGAG 240  
 DB 181 GGGCAAGGAGAACTCGAGCTCAGATGCTCAGCAGCACTTTGTTAGTAAATCGGAG 240  
 QY 241 GTGAAGATGAGATATTCATCAGAGTTTTCAGTCACTTCTTGTGCAATTAACAGAACTG 300  
 DB 241 GTGAAGATGAGATATTCATCAGAGTTTTCAGTCACTTCTTGTGCAATTAACAGAACTG 300  
 QY 301 ACACCTGAATATGCTTGAAGCTGCAAGCGGCTTTTGGAGAAAGTCTTATGATTTCC 360  
 DB 301 ACACCTGAATATGCTTGAAGCTGCAAGCGGCTTTTGGAGAAAGTCTTATGATTTCC 360  
 QY 361 TCACAGGTTTACAGATTCCTGTGGCAATTTCTACCAAGCAAGCTAAGACAGTAGACT 420  
 DB 361 TCACAGGTTTACAGATTCCTGTGGCAATTTCTACCAAGCAAGCTAAGACAGTAGACT 420  
 QY 421 TTGTGAATGATACAGAGAACTCCACACAGTGTAACTCTCTGGGTTGCTGATAAACTA 480  
 DB 421 TTGTGAATGATACAGAGAACTCCACACAGTGTAACTCTCTGGGTTGCTGATAAACTA 480  
 QY 481 AAGCCTGGAATATTTCAACAGAGCTGTACATCTGAGAGAGCCAGGAATCGCTCTT 540  
 DB 481 AAGCCTGGAATATTTCAACAGAGCTGTACATCTGAGAGAGCCAGGAATCGCTCTT 540  
 QY 541 CCTCTGTGTACTGCAAGCTCGCTTTTACAGCCCTTACTGCTTCACTCTATTCCTCAAT 600  
 DB 541 CCTCTGTGTACTGCAAGCTCGCTTTTACAGCCCTTACTGCTTCACTCTATTCCTCAAT 600  
 QY 601 GCAACTCTCTGTGACCCCGCATGCTGTGTCGTCCTCCCTGCTGTGAGCAGNTGTA 660  
 DB 601 GCAACTCTCTGTGACCCCGCATGCTGTGTCGTCCTCCCTGCTGTGAGCAGNTGTA 660  
 QY 661 CTAATAAATGCGGCAATTTTCACTGTAAATAAAAAAAAAAAAAA 706  
 DB 661 CTAATAAATGCGGCAATTTTCACTGTAAATAAAAAAAAAAAAAA 706

RESULT 3  
 ID AAA74934  
 XX AAA74934 standard; DNA; 599 BP.  
 AC AAA74934;  
 XX  
 DT 02-JAN-2001 (first entry)  
 DE Nucleotide sequence encoding a human serpin polypeptide.  
 XX Human; serpin; serine protease inhibitor; autoimmune disease; neoplasm;  
 KW rheumatoid arthritis; hyperproliferative disorder; cerebral ischemia;

KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
 KW ocular disorder; corneal infection; epithelial cell proliferation;  
 KW skin aging; sunburn; wound healing; organ transplantation; chemotaxis;  
 KW food additive; food preservative; ss.  
 XX Homo sapiens.  
 XX OS  
 XX FH Key Location/Qualifiers  
 XX CDS 1..597  
 FT /tag= a  
 FT /transl\_except= (pos: 583..585, aa: Xaa)  
 FT /product= "serpin"  
 FT /note= "Xaa is any L-amino acid; no termination codon  
 given"

WO200052160-A1.  
 08-SEP-2000.  
 29-FEB-2000; 2000WO-US05082.  
 01-MAR-1999; 99US-0122276.  
 12-MAR-1999; 99US-0124094.  
 18-AUG-1999; 99US-0149452.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Ruben SM, Ni J;  
 WPI; 2000-579284/54.  
 P-PSDB; AAB08783.  
 Novel human serine protease inhibitor superfamily of proteins useful  
 for diagnosing, treating, preventing and/or prognosing disorders  
 related to the Serpin polypeptides -  
 Claim 1; Page 12-13; 215pp; English.

The present sequence encodes a human serpin polypeptide. Serpin is  
 a member of the serine protease inhibitor superfamily of polypeptides.  
 The serpin polynucleotides and polypeptides are used to prevent, treat  
 or ameliorate a medical condition in e.g. humans, mice, rabbits, goats,  
 horses, cats, dogs, chickens or sheep. Disorders which are diagnosed or  
 treated include autoimmune diseases e.g. rheumatoid arthritis  
 hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 e.g. cerebral ischemia, angiogenesis, nervous system disorders,  
 e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
 fungi and ocular disorders e.g. corneal infection. The polypeptides can  
 also be used to aid wound healing and epithelial cell proliferation,  
 to prevent skin aging due to sunburn, to maintain organs before  
 transplantation, for supporting cell culture of primary tissues, to  
 regenerate tissues and in chemotaxis. The polypeptides can also be used  
 as a food additive or preservative to increase or decrease storage  
 capabilities.

Sequence 599 BP; 171 A; 141 C; 122 G; 164 T; 1 other;  
 Query Match 83.0%; Score 586; DB 21; Length 599;  
 Best Local Similarity 99.8%; Pred. No. 8.7e-164;  
 Matches 598; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 52 ATGATGCTCTATCAGAGCAAAATGGCAGATTTTGCATTAACTTTTGAAGAACTAGGG 111  
 DB 1 ATGATGCTCTATCAGAGCAAAATGGCAGATTTTGCATTAACTTTTGAAGAACTAGGG 60  
 QY 112 GAAACAACTCAAACTTATTTTTCATGAGCATATCATCAGCTTGGCCATG 170  
 DB 61 GAAACAACTCAAACTTATTTTTCATGAGCATATCATCAGCTTGGCCATG 120  
 QY 171 GTTTTCATGGGCAAGGGAACACTGCAGCTCAGATGCTCAGGCACTTTTGTAGT 230



DE Nucleotide sequence encoding a human serpin polypeptide.

XX Human; serpin; serine protease inhibitor; autoimmune disease; neoplasm;  
 XX rheumatoid arthritis; hyperproliferative disorder; cerebral ischemia;  
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
 KW ocular disorder; corneal infection; epithelial cell proliferation;  
 KW skin aging; sunburn; wound healing; organ transplantation; chemotaxis;  
 KW food additive; food preservative; ss.

OS Homo sapiens.

PH Key Location/Qualifiers  
 FT CDS 1..606  
 FT /\*tag= a  
 FT /transl\_except= (pos: 91..93, aa: Xaa)  
 FT /transl\_except= (pos: 94..96, aa: Xaa)  
 FT /transl\_except= (pos: 97..99, aa: Xaa)  
 FT /product= "serpin"  
 FT /note= "Xaa is any L-amino acid; no termination codon given"

XX WO200052160-A1.  
 XX 08-SEP-2000.  
 XX 29-FEB-2000; 2000WO-US05082.  
 XX 01-MAR-1999; 99US-0122276.  
 XX 12-MAR-1999; 99US-0124094.  
 XX 18-AUG-1999; 99US-0149452.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Ruben SM, Ni J;  
 XX WPI; 2000-579284/54.  
 XX P-PSDB; AAB08784.

PT Novel human serine protease inhibitor superfamily of proteins useful  
 PT for diagnosing, treating, preventing and/or prognosing disorders  
 PT related to the Serpin polypeptides

XX Claim 1; Page 13; 215pp; English.

CC The present sequence encodes a human serpin polypeptide. Serpin is  
 CC a member of the serine protease inhibitor superfamily of polypeptides.  
 CC The serpin polynucleotides and polypeptides are used to prevent, treat  
 CC or ameliorate a medical conditions in e.g. humans, mice, rabbits, goats,  
 CC horses, cats, dogs, chickens or sheep. Disorders which are diagnosed or  
 CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischemia, angioneurosis, nervous system disorders.  
 CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
 CC fungi and ocular disorders e.g. corneal infection. The polypeptides can  
 CC also be used to aid wound healing and epithelial cell proliferation,  
 CC to prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities.

XX Sequence 608 BP; 171 A; 141 C; 122 G; 164 T; 10 other;  
 SQ

Query Match 81.7%; Score 577; DB 21; Length 608;  
 Best Local Similarity 98.4%; Pred. No. 4.1e-161;  
 Matches 598; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 52 ATGGATGCTCTATCAGAGCAATGGACATTTGCATTAAACCTTTTGAAAAGCTAGGG 111  
 DB 1 ATGGATGCTCTATCAGAGCAATGGACATTTGCATTAAACCTTTTGAAAAGCTAGGG 60

QY 112 GAAGCAACTCAACAACTTATTTTTC-----CCCATGAGCATATCATCAGCC 161  
 DB 61 GAAACAACCTCAACAACTTATTTTTCNNNNNNNNNNCCCATGAGCATATCATCAGCC 120  
 QY 162 TTGGCCATGGTTTTCATGGGGCAAGGAAACACTGCGAGCTCAGATGCTCAGGCACTT 221  
 DB 121 TTGGCCATGGTTTTCATGGGGCAAGGAAACACTGCGAGCTCAGATGCTCAGGCACTT 180  
 QY 222 TGTTTTAGTAATAATCGGAGGTGAAGATGAGAGATATTCATCGAGGTTTTCAGTCACTTCTT 281  
 DB 181 TGTTTTAGTAATAATCGGAGGTGAAGATGAGAGATATTCATCGAGGTTTTCAGTCACTTCTT 240  
 QY 282 GTTGCAATTAACAGAACTGACACTGAATATGCTGTAGAACTGCCAAGGGCTCTTTTGA 341  
 DB 241 GTTGCAATTAACAGAACTGACACTGAATATGCTGTAGAACTGCCAAGGGCTCTTTTGA 300  
 QY 342 GAAAGTCTTATGATTTCTTCCACAGGTTTACAGATTCCTGTGGCAAAATTCACCAAGCA 401  
 DB 301 GAAAGTCTTATGATTTCTTCCACAGGTTTACAGATTCCTGTGGCAAAATTCACCAAGCA 360  
 QY 402 ACGATAAACACAGCTAGACTTTTGTGAATGATACAGAGAGTCCACACAGGTGTAACCTCC 461  
 DB 361 ACGATAAACACAGCTAGACTTTTGTGAATGATACAGAGAGTCCACACAGGTGTAACCTCC 420  
 QY 462 TGGTTGCTGATATAAACTAAAGCCTGGAATAATTATTCAAACAAAGCTGTACATCTGGAG 521  
 DB 421 TGGTTGCTGATATAAACTAAAGCCTGGAATAATTATTCAAACAAAGCTGTACATCTGGAG 480  
 QY 522 GAGCCAGGAATCGCTCTTCTCTTGTACTGCAAGCCCTGCTTTCACAGCCCTACTG 581  
 DB 481 GAGCCAGGAATCGCTCTTCTCTTGTACTGCAAGCCCTGCTTTCACAGCCCTACTG 540  
 QY 582 GTTCACCTTATTTCCCAATGCACTCTCCTGTGACCCCGCATGGCATGTGTGNCCTCCC 641  
 DB 541 GTTCACCTTATTTCCCAATGCACTCTCCTGTGACCCCGCATGGCATGTGTGNCCTCCC 600

QY 642 TCCCTGTG 649  
 DB 601 TCCCTGTG 608

RESULT 6  
 AAL42748  
 ID AAL42748 standard; cDNA; 608 BP.  
 XX AC AAL42748;  
 XX AC AAL42748;  
 XX DT 19-JUL-2002 (first entry)  
 XX DE Human serine protease inhibitor (serpin) coding sequence 2.  
 XX DE Human; serine protease inhibitors; serpin; blood coagulation;  
 KW proteinase-activated physiological process; gene therapy;  
 KW antitense therapy; cancer; inflammation; neurological disease;  
 KW immune disease; gene; ss.  
 XX OS Homo sapiens.  
 XX PH Key Location/Qualifiers  
 FT CDS 1..606  
 FT /\*tag= a  
 FT /product= "Human serpin protein 2"  
 FT /transl\_except= (pos: 91..93, aa: Xaa)  
 FT /transl\_except= (pos: 94..96, aa: Xaa)  
 FT /transl\_except= (pos: 97..99, aa: Xaa)  
 FT /transl\_except= (pos: 992..594, aa: Xaa)  
 FT /note= "Xaa is unknown"

XX WO200155390-A1.  
 XX 02-AUG-2001.  
 XX 26-JAN-2001; 2001WO-US02484.

XX	28-JAN-2000; 2000US-0178769.	582	GTTCACCTATTCCCAATGCAACTCTCTCTGTGACCCCGCATGCGATGTGTGTCCTCC	QY	582	GTTCACCTATTCCCAATGCAACTCTCTCTGTGACCCCGCATGCGATGTGTGTCCTCC
PR	29-FEB-2000; 2000WO-US05082.	541	GTTCACCTATTCCCAATGCAACTCTCTCTGTGACCCCGCATGCGATGTGTGTCCTCC	Db	541	GTTCACCTATTCCCAATGCAACTCTCTCTGTGACCCCGCATGCGATGTGTGTCCTCC
XX	(HUMA-) HUMAN GENOME SCI INC.	642	TCCTCTGTG 649	QY	642	TCCTCTGTG 649
XX	Ni J, Ruben SM, Shi Y;	601	TCCTCTGTG 608	Db	601	TCCTCTGTG 608
PI	WPI; 2001-457730/49.					
XX	P-PSDB; AA014922.					
DR	Isolated serine protease inhibitor polypeptide for treating, preventing					
XX	and/or prognosing cancers, inflammation, immune disorders and					
PT	neurological diseases and also for testing and detection e.g. diagnosis					
PT						
PT						
XX	Claim 1; Page 15; 288pp; English.					
PS	The invention comprises the amino acid and coding sequences of a group of					
XX	novel human serine protease inhibitors (serpins). Serpins are a					
CC	superfamily of single chain proteins that contain a conserved structure					
CC	of approximately 370-420 amino acids and generally range between 50 and					
CC	100 kDa in molecular mass. The majority of serpins function as protease					
CC	inhibitors and so are involved in the regulation of several proteinase-					
CC	activated physiological processes (e.g. blood coagulation). The Serpin					
CC	DNA and amino acid sequences of the invention are useful in the treatment					
CC	(e.g. gene therapy, antisense therapy) and diagnosis of conditions such					
CC	as cancer, inflammation, neurological diseases and immune diseases. The					
CC	present cDNA sequence represents a human serpin coding sequence of the					
CC	invention.					
XX	Sequence 608 BP; 171 A; 141 C; 122 G; 164 T; 10 other;					
SQ						
	Query Match 81.7%; Score 577; DB 23; Length 608;					
	Best Local Similarity 98.4%; Pred. No. 4.1e-161;					
	Matches 598; Conservative 0; Mismatches 0; Indels 10; Gaps 1;					
QY	52 ATGGATGCTCTATCAGCAAGCAATGCGACATTTGCAATTAACCTTTTGAAGAGCTAGGG 111					
Db	1 ATGGATGCTCTATCAGCAAGCAATGCGACATTTGCAATTAACCTTTTGAAGAGCTAGGG 60					
QY	112 GAAACAACTCAACAACTTATTTTTC-----CCATGAGCATATCATCAGCC 161					
Db	61 GAAACAACTCAACAACTTATTTTTCNNNNNNNNCCATGAGCATATCATCAGCC 120					
QY	162 TTGGCCATGTTTTCATGGGGCAAGGAAACACTGCAGCTCAGATGCTCAGGCACCT 221					
Db	121 TTGGCCATGTTTTCATGGGGCAAGGAAACACTGCAGCTCAGATGCTCAGGCACCT 180					
QY	222 TGTTTTAGTAAATCGGAGGTGAAGATGGAGATATTTCATCGAGGTTTTCAGTCACCTT 281					
Db	181 TGTTTTAGTAAATCGGAGGTGAAGATGGAGATATTTCATCGAGGTTTTCAGTCACCTT 240					
QY	282 GTTGCAATTAAAGAACTGACATGTAATATGCTTAGAATCTGCAACGGGCTCTTGA 341					
Db	241 GTTGCAATTAAAGAACTGACATGTAATATGCTTAGAATCTGCAACGGGCTCTTGA 300					
QY	342 GAAAGTCTTATGATTTCTTCCACAGGTTTACAGATTCCTGTGCAAAATTTTACCAAGCA 401					
Db	301 GAAAGTCTTATGATTTCTTCCACAGGTTTACAGATTCCTGTGCAAAATTTTACCAAGCA 360					
QY	402 ACGATAAAACAGCTAGACTTTTGTGAATGATACAGAGAAGTCCAAACAGCTGTAAACTCC 461					
Db	361 ACGATAAAACAGCTAGACTTTTGTGAATGATACAGAGAAGTCCAAACAGCTGTAAACTCC 420					
QY	462 TGGTTCTGATTAACCTGAAATTTTAAACAGAGCTGTACATCTGGAG 521					
Db	421 TGGTTCTGATTAACCTGAAATTTTAAACAGAGCTGTACATCTGGAG 480					
QY	522 GAGCCAGGATCGGCTCTTCTCTTTGTTACTGCAAGCCTGCTTTTACAGAGCCCTACTG 581					
Db	481 GAGCCAGGATCGGCTCTTCTCTTTGTTACTGCAAGCCTGCTTTTACAGAGCCCTACTG 540					
QY	582 GTTCACCTATTCCCAATGCAACTCTCTCTGTGACCCCGCATGCGATGTGTGTCCTCC					
Db	541 GTTCACCTATTCCCAATGCAACTCTCTCTGTGACCCCGCATGCGATGTGTGTCCTCC					
QY	642 TCCTCTGTG 649					
Db	601 TCCTCTGTG 608					
RESULT 7						
ABK94967						
ID	ABK94967 standard; CDNA; 943 BP.					
XX	AC					
XX	ABK94967;					
XX	30-AUG-2002 (first entry)					
DT						
XX	Human novel polynucleotide #78.					
DE						
XX	Human; gene; ss; inflammatory condition; shock; sepsis; immune response;					
XX	cancer; wound healing; central nervous system disease; haematopoiesis;					
XX	peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;					
XX	myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;					
XX	cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;					
XX	bone degenerative disorder; periodontal disease; reperfusion injury;					
XX	lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;					
XX	allergic condition; thrombolysis; thrombosis; coagulation disorder;					
XX	fungal infection.					
OS	Homo sapiens.					
XX	WO200244340-A2.					
XX	06-JUN-2002.					
PD						
XX	30-NOV-2001; 2001WO-US47004.					
PF						
XX	30-NOV-2000; 2000US-0028952.					
PR						
XX	(HYSE-) HYSEQ INC.					
XX	Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;					
PI	Yamazaki V, Ujwal ML, Dmanac RT;					
PI	WPI; 2002-508509/54.					
XX	P-PSDB; ABG66743.					
DR						
XX	Novel nucleic acids and polypeptides for diagnosis, treatment of					
XX	inflammatory, autoimmune, nervous system, myeloid or lymphoid cell					
XX	disorders, cancer and promoting wound healing -					
XX	Claim 1; Page 524-525; 672pp; English.					
XX	The invention relates to human novel polynucleotides and associated					
XX	polypeptides. The polynucleotides and polypeptides are useful for					
XX	treating inflammatory conditions such as arthritis, nephritis, Crohn's					
XX	disease, ischaemia-reperfusion injury, shock, sepsis, immune responses					
XX	and cancer and for promoting wound healing. The sequences are used to					
XX	induce the proliferation of neural cells and regeneration of nerve and					
XX	brain tissue, and are useful for the treatment of central and peripheral					
XX	nervous system diseases and neuropathies, such as Alzheimer's disease,					
XX	Parkinson's disease, Huntington's disease and amyotrophic lateral					
XX	sclerosis. The sequences are involved in chemotactic or chemokinetic					
XX	activity, regulation of haematopoiesis, treatment of myeloid or lymphoid					
XX	cell disorders and platelet disorders such as thrombocytopenia,					
XX	regeneration of bone, cartilage, tendon, ligament and/or nerve tissue					
XX	growth, tissue repair, healing of burns, incisions, ulcers, treatment of					
XX	osteoporosis, osteoarthritis, bone degenerative disorders and periodontal					
XX	disease. The sequences of the invention are also useful for gut					
XX	protection or regeneration and treatment of lung or liver fibrosis,					
XX	reflux injury in various tissues, immune deficiencies and disorders					
XX	including severe combined immunodeficiency (SCID), bacterial or fungal					
XX	infections, autoimmune disorders e.g. multiple sclerosis and myasthenia					



CC gravis, allergic conditions such as asthma, thrombolysis or thrombosis  
 CC and coagulation disorders. Sequences ABK94890-ABK94982 represent human  
 CC novel polynucleotides of the invention.

XX SQ Sequence 943 BP; 271 A; 214 C; 186 G; 272 T; 0 other;

Query Match 65.6%; Score 462.8; DB 24; Length 943;

Best Local Similarity 98.4%; Pred. No. 4.2e-127;

Matches 478; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 9 CTTTCGCTCTGGGACGTCGGGAGAACCTGGGGCTCAGCGTATGATGCTCTATCAGA 68  
 DB 146 CTTTCGCTCTGGGACGTCGGGAGAACCTGGGGCTCAGCGTATGATGCTCTATCAGA 205

QY 69 AGCAATGGCACATTTGCATTAAACCTTTTGAAGAGCTAGGGCAAAACAACCAACAA 128  
 DB 206 AGCAATGGCACATTTGCATTAAACCTTTTGAAGAGCTAGGGCAAAACAACCAACAA 265

QY 129 CTTATTTTTCCTCCCA-TGAGCATATCATCAGCCTTGGCCATGGTTTTCATGGGGCAAA 187  
 DB 266 CTTATTTTTCCTCCCACTGAGCATATCATCAGCCTTGGCCATGGTTTTCATGGGGCAAA 325

QY 188 GGGAAACACTGACGCTCAGATGCTCAGGCACCTTTCTTTAGTAAATCGAGGTCAAGA 247  
 DB 326 GGGAAACACTGACGCTCAGATGCTCAGGCACCTTTCTTTAGTAAATCGAGGTCAAGA 385

QY 248 TGAGATATTTCATCGAGGTTTTCAGTCACCTTCTTGTGCAATTAACAGAACTGACACTGA 307  
 DB 386 TGAGATATTTCATCGAGGTTTTCAGTCACCTTCTTGTGCAATTAACAGAACTGACACTGA 445

QY 308 ATATGCTTAGAAGTGCACGCGGCTCTTTGAGAAAGTCTTATGATTTCTCACAGG 367  
 DB 446 ATATGCTTAGAAGTGCACGCGGCTCTTTGAGAAAGTCTTATGATTTCTCACAGG 505

QY 368 TTTTACAGATTCTCTGGGCAATTTTACCAAGCAAGTATTAACACACTAGACTTTGTGAA 427  
 DB 506 TTTTACAGATTCTCTGGGCAATTTTACCAAGCAAGTATTAACACACTAGACTTTGTGAA 565

QY 428 TGATCAGAGAAAGTCCACACACGCTGTAAACTCTGGGTTGCTGATAAAACTAAAGCCTG 487  
 DB 566 TGATCAGAGAAAGTCCACACACGCTGTAAACTCTGGGTTGCTGATAAAACTAAAGTGA 625

QY 488 GAAAT 493

DB 626 AATAT 631

RESULT 8

AAF16065

ID AAF16065 standard; cDNA; 1624 BP.

AC AAF16065;

XX 13-MAR-2001 (first entry)

DT Human prostate cancer antigen nucleotide sequence SEQ ID NO: 500.

DE Human prostate cancer antigen nucleotide sequence SEQ ID NO: 500.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
 KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;  
 KW vulnary; gastrointestinal; nephrotropic; anti-infective; gynaecological;  
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
 KW wound; infectious disease; ss.

XX Homo sapiens.

OS WO200055174-A1.

PN 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05988.

PF 12-MAR-1999; 99US-0124270.

XX

PR

XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.

XX PI Rosen CA, Ruben SM;

XX WIPI; 2000-587513/55.

DR P-PSDB; AAB56862.

XX PT Prostate cancer associated gene sequences, referred to as prostate

PT cancer antigens, useful for treatment, prevention, and diagnosis of  
 PT disorders such as prostate cancer -

XX Claim 1; Page 967; 2338pp; English.

XX AAF15566 to AAF15505 encode the human prostate cancer associated

CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

CC The prostate cancer antigens can have neuroprotective, cytosolic,

CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,

CC nephrotropic, anti-infective, gynaecological and antibacterial activities,

CC and can be used in gene therapy. The prostate cancer antigen

CC polynucleotides may be used for detection of prostate cancer, chromosome

CC identification, as chromosome markers, and for numerous other diagnostic

CC or research purposes. The prostate cancer antigens may be used to treat

CC disorders such as neural, immune, muscular, reproductive,

CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to

CC AAB57303 represent sequences used in the exemplification of the present

CC invention.

XX Sequence 1624 BP; 414 A; 421 C; 430 G; 356 T; 3 other;

SQ Query Match 35.6%; Score 251.2; DB 21; Length 1624;

Best Local Similarity 73.6%; Pred. No. 3.9e-64;

Matches 346; Conservative 1; Mismatches 119; Indels 4; Gaps 2;

QY 15 TCCTGGGAGCTGCGGAGAACTGGGCTCACCGCTCATGGATGCTCTATCAGAAAGCAAA 74

DB 277 TTCTGCTCGCTCGCTCCCGCTCTGGAGTCTGCCATCATGATGTTCTCGCAGAGCAAA 336

QY 75 TGGCAGATTGATTAACCTTTTGAAGAGTAGGGGAAACAACACTCAAACTATTATT 134

DB 337 TGGCAGCTTTGCTTTAAACCTTTTGAAGAGTAGGGGAAACAACACTCGAAGATGTGT 396

QY 135 TTTTTC-CCCATGAGCATATCATCAGCCTTGCCCATGTTTTCATGGGGCAAGGAAA 193

DB 397 TTTCTACCCATGAGCATGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 456

QY 194 CACTGAGCTCAGATGCTCTAGGCACTTTGTTTTAGTAAATCGGAGGTGAAGATGGAGA 253

DB 457 CACCGCTGCACAGATGGCCAGATACCTTTTCAATAAAGTGCGGT---GGTGGAGA 513

QY 254 TATTCATCGAGTTTTCAGTCACTTCTTGTGCAATTAACAGAACTGACACTGAATATGT 313

DB 514 CATCCACAGGGCTTCCAGTCTCTTCTCACCAAGTGAACAGACTGCGCAGTACTT 573

QY 314 GCTTAGAAGTCCCAACGGGCTCTTTGGAGAAAAGTCTTATGATTTTCTCACAGGTTTAC 373

DB 574 GCTTAGGATGCGCAACAGGCTCTTTGGGAAAAGTCTTGTGATTTTCTCTCATCTTTAG 633

QY 374 AGATTCTCTGTGGCAATTTCTACCAAGCAACGATAAACAGCTAGACTTTGTGAATGATAC 433

DB 634 AGATTCTCTGGCAAAATTTCTACCAAGCAGAGATGGAGGCTTGACTTTATCAGGCGCT 693

QY 434 AGAAGTCCCAACAGCTGTAAACTCTCTGGTTGCTGTGATAAACTAAAG 483

DB 694 AGAAGTCCCAAGAAACACATAAACCTCTGGTAGCTGAGTGAAGAGAGAG 743

RESULT 9

AAS06075

ID AAS06075 standard; DNA; 695 BP.

XX

AC AAS06075;  
XX 12-SEP-2001 (first entry)  
XX Angiotensin converting enzyme (ACEV) splice variant DNA #75.  
DE XX  
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;  
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;  
KW platelet-derived endothelial cell growth factor; cardiovascular disease;  
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C; ds;  
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;  
KW myocardial infarction; coronary arterial thrombosis; renal disease;  
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;  
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;  
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;  
KW vascular disorder; asbestosis.  
XX  
OS Homo sapiens.  
XX  
XX WO200136632-A2.  
XX  
XX 25-MAY-2001.  
XX  
XX 17-NOV-2000; 2000WO-IL00766.  
XX  
XX 17-NOV-1999; 99IL-0132978.  
XX  
XX 10-DEC-1999; 99IL-0133455.  
XX  
XX (COMP-) COMPUGEN LTD.  
XX  
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;  
XX WPI: 2001-336004/35.  
XX P-PSDB; AAU02975.  
XX  
XX Novel alternative splicing variants e.g. variant of angiotensin  
XX converting enzyme (ACEV), useful in identifying candidate compounds  
XX capable of binding to the variant and to detect anti-variant antibodies  
XX  
XX Claim 1; Page 367; 519pp; English.  
XX  
XX The sequence represents a DNA encoding an angiotensin converting enzyme  
XX splice variant (ACEV) polypeptide. The polypeptides of the invention  
XX include variants of granulocyte colony stimulating factor receptor,  
XX glucagon, interleukin 6, platelet-derived endothelial cell growth factor,  
XX cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P53, and  
XX vasoactive intestinal polypeptide receptor 2. The polypeptides and their  
XX associated nucleic acids are useful for identification of variant  
XX sequences and detection of candidate compounds capable of binding the  
XX molecules. The sequences of the invention can be used in the treatment  
XX and diagnosis of various disorders including cardiovascular diseases such  
XX as arteriosclerosis, myocardial infarction and coronary arterial  
XX thrombosis, renal diseases such as diabetic nephropathy, muscular  
XX diseases such as hypertrophy, immune disorders such as immune complex  
XX nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic  
XX pulmonary granulomatous diseases such as asbestosis and vascular  
XX pathologies involving an endothelial abnormality such as deep vein  
XX thrombosis.  
XX  
XX Sequence 695 BP; 184 A; 158 C; 154 G; 199 T; 0 other;  
XX

Query Match 35.5%; Score 250.4; DB 22; Length 695;  
Best Local Similarity 75.4%; Pred. No. 4.5e-64;  
Matches 338; Conservative 0; Mismatches 106; Indels 4; Gaps 2;  
QY 37 CTGGGGCTCACCCTGATGCTCTATCAGAGCAATGGCAGATTCGATTAACCTT 96  
DB 60 CTGGGCTCGCCATCATGATGTTCTCGCAGAGCAATGGCAGCTTTGCCCTTAAACCTT 119  
QY 97 TTGAAAGCTAGGGGAAACAACTCAAACTATTTTTC-CCCATGAGCATACA 155  
DB 120 TTGAAAGCTAGGGGAAACAACTCAAACTATTTTTC-CCCATGAGCATACA 179

QY 156 TCAGCCTTGGCCATGGTTTTCATGGGGCAAAAGGAAACACTGCAGCTCAGATGTCTCAG 215  
DB 180 TGTGCCCTGGCCATGGTCTACATGGGGCAAAAGGAAACACCGCTGCACAGATGCCCCAG 239  
QY 216 GCATTTCTTTTAGTAAATCGAGGTGAAGATGAGATATTCATCGAGGTTTTCAGTCA 275  
DB 240 ATACTTTCTTCAATAAAAGTGGCGGT--GGTGGAGACATCCACAGGGGTTCCAGTCT 296  
QY 276 CTTCTTGTGCAATTAACAGAACTGACACTGAATATGCTTAGAACTGCCAACGGGCTC 335  
DB 297 CTTCTCACCGAAGTGAAACAGACTGGCAGCAGTACTTCTTAGATGGCCCAACAGGCTC 356  
QY 336 TTTGAGAGAAAGTCTTATGATTTCTCTACAGGTTTACAGATTTCTGTGGCAAAATTTTAC 395  
DB 357 TTTGGGAAAAGTCTTGTGATTTCTCTCATCTTTTAGAGATTCCTGCCAAAATTTTAC 416  
QY 396 CAAGCAACGATATAAACAGCTAGACTTTGTGAATATGATACAGAGAGTCCACACAGTGA 455  
DB 417 CAAGCAGAGATGGAGGAGCTTGTGATTTATAGCGCGGTAGAGAGTCCAGAAACACATA 476  
QY 456 AACTCTCGGTTGCTGATAAAACTAAAG 483  
DB 477 AACACCTGGTAGCTGAAAAGACAGAAAG 504  
RESULT 10  
AAS06074  
ID AAS06074 standard; DNA; 868 BP.  
XX  
XX AAS06074;  
XX AC  
XX 12-SEP-2001 (first entry)  
XX DT  
XX DE Angiotensin converting enzyme (ACEV) splice variant DNA #74.  
XX  
XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;  
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;  
KW platelet-derived endothelial cell growth factor; cardiovascular disease;  
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C; ds;  
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;  
KW myocardial infarction; coronary arterial thrombosis; renal disease;  
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;  
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;  
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;  
KW vascular disorder; asbestosis.  
XX  
XX Homo sapiens.  
XX  
XX WO200136632-A2.  
XX  
XX 25-MAY-2001.  
XX  
XX 17-NOV-2000; 2000WO-IL00766.  
XX  
XX 17-NOV-1999; 99IL-0132978.  
XX  
XX 10-DEC-1999; 99IL-0133455.  
XX  
XX (COMP-) COMPUGEN LTD.  
XX  
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;  
XX WPI: 2001-336004/35.  
XX P-PSDB; AAU02974.  
XX  
XX Novel alternative splicing variants e.g. variant of angiotensin  
XX converting enzyme (ACEV), useful in identifying candidate compounds  
XX capable of binding to the variant and to detect anti-variant antibodies  
XX  
XX Claim 1; Page 366; 519pp; English.  
XX  
XX The sequence represents a DNA encoding an angiotensin converting enzyme

splice variant (ACEV) polypeptide. The polypeptides of the invention include variants of granulocyte colony stimulating factor receptor, glucagon, interleukin 6, platelet-derived endothelial cell growth factor, cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P33, and vasoactive intestinal polypeptide receptor 2. The polypeptides and their associated nucleic acids are useful for identification of variant sequences and detection of candidate compounds capable of binding the molecules. The sequences of the invention can be used in the treatment and diagnosis of various disorders including cardiovascular diseases such as arteriosclerosis, myocardial infarction and coronary arterial thrombosis, renal diseases such as diabetic nephropathy, muscular diseases such as hypertrophy, immune disorders such as immune complex nephritis, multiple sclerosis, cancer, sarcoidosis, nonarclotoidic pulmonary granulomatous diseases such as asbestosis and vascular pathologies involving an endothelial abnormality such as deep vein thrombosis.

Sequence 868 BP; 214 A; 212 C; 227 G; 215 T; 0 other;

Query Match 35.5%; Score 250.4; DB 22; Length 868;  
 Best Local Similarity 75.4%; Pred. No. 5e-64;  
 Matches 338; Conservative 0; Mismatches 106; Indels 4.

Qy	37	CTGGGGCTCACCGT	CATGGATGCTCTAT	CGAAGCAAAATGG	CACATTTGCATTAACCTT	96
Db	60	CTTGGGTC	TGCGCATCATGGATGTTCTCG	CAGAAGCAAAATGG	CACCTTTGCCTTAAACCTT	119
Qy	97	TTGAAAAGCT	TAGGGGAAAACA	CTCAAAACACTTATTTTTC	-CCCATGAGCATATCA	155
Db	120	TTGAAAACGCT	GGGTAAAGACAC	CTCGAAGAATGTGTTTTCT	CACCATGAGCATGTCC	179
Qy	156	TCAGCCTTGGCCAT	TGGTTTTTCATGGGGCAAAAGG	GAACA	CACTGTCAGCTCAGATGTCTCAG	215
Db	180	TGTGCCCTGGCCAT	GGTCTACATGGGGCAAAAGG	GAACA	CCGCTGCACAGATGCCCCAG	239
Qy	216	GCATTTGTTTTT	TAGTAAAACTCGGAGGTGAAG	ATCGAGATATTCATCGAGGTTTT	TCAGTCA	275
Db	240	ATACTTTCTTT	CAATAAAGTGGCGGT---	GGTGAGACATCCACAGGGCTTC	CAGTCT	296
Qy	276	CTTCTTGTTGCA	TTAAACAGAACTTGACA	CTGAATATGTGCTTAGAACTGCCAA	CGGGTCT	335
Db	297	CTTCTCACCGA	AGTGAACAAGACTTGGCACG	CAGTACTTGCTTAGGATGCCCA	CAGGGTCT	356
Qy	336	TTTGGAGAAAAGT	CTTATGATTTCTCTACAGGTTTT	CAGATTTCTGTGGCAAAATTC	TCTAC	395
Db	357	TTTGGGAAAAGT	CTTGTAATTTCTCTCATCTTTTAGAGAT	TCTGTGCAAAAAATTC	TCTAC	416
Qy	396	CAAGCAACGAT	AAAAACAGCTAGACTTTGTG	TAATGATACAGAGAAAGTCC	CACAACAGCTGA	455
Db	417	CAAGCAGAT	TGGAGGAGCTTGACTTTAT	CAGCGCGGTAGAGAAAGTCC	AGAAAACACATA	476
Qy	456	AATCTCTGGG	TTGCTGATAAAACT	TAAG	483	
Db	477	AAACCTTGGG	TAGCTGAAAAACAGAGA	504		

**RESULT 11**

AAS06073  
ID AAS06073 standard; DNA; 1302 BP.

AC AAS06073;

DT 12-SEP-2001 (first entry)

DE Angiotensin converting enzyme (ACEV) splice variant DNA #73.

Angiotensin converting enzyme splice variant; ACEV; interleukin 6;  
granulocyte colony stimulating factor receptor; glucagon; hypertrophy;  
platelet-derived endothelial cell growth factor; cardiovascular disease;  
tumour antigen P53; cyclin-dependent kinase inhibitor 1C; ds;  
cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C; ds;  
vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;  
myocardial infarction; coronary arterial thrombosis; renal disease;  
diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;

multiple sclerosis; immune complex nephritis; deep vein thrombosis; nonarcolotic pulmonary granulomatous disease; endothelial abnormality; vascular disorder; asbestosis.

**Homo sapiens.**

WO200136632-A2

25-MAY-2001.

17-NOV-2000:

17-NOV-1999: . 99IL-0132978.

10-DEC-1999; 99IL-0133455.

(COMP-) COMPUGEN LTD.

Levine Z, David A, Azar I, Khosravi R, Bernstein J;

WPI; 2001-336004/35.

1. **Introduction**  
 2. **Background**  
 3. **Methodology**  
 4. **Results**  
 5. **Discussion**  
 6. **Conclusion**  
 7. **References**  
 8. **Appendix**  
 9. **Figure 1**  
 10. **Figure 2**  
 11. **Figure 3**  
 12. **Figure 4**  
 13. **Figure 5**  
 14. **Figure 6**  
 15. **Figure 7**  
 16. **Figure 8**  
 17. **Figure 9**  
 18. **Figure 10**  
 19. **Figure 11**  
 20. **Figure 12**  
 21. **Figure 13**  
 22. **Figure 14**  
 23. **Figure 15**  
 24. **Figure 16**  
 25. **Figure 17**  
 26. **Figure 18**  
 27. **Figure 19**  
 28. **Figure 20**  
 29. **Figure 21**  
 30. **Figure 22**  
 31. **Figure 23**  
 32. **Figure 24**  
 33. **Figure 25**  
 34. **Figure 26**  
 35. **Figure 27**  
 36. **Figure 28**  
 37. **Figure 29**  
 38. **Figure 30**  
 39. **Figure 31**  
 40. **Figure 32**  
 41. **Figure 33**  
 42. **Figure 34**  
 43. **Figure 35**  
 44. **Figure 36**  
 45. **Figure 37**  
 46. **Figure 38**  
 47. **Figure 39**  
 48. **Figure 40**  
 49. **Figure 41**  
 50. **Figure 42**  
 51. **Figure 43**  
 52. **Figure 44**  
 53. **Figure 45**  
 54. **Figure 46**  
 55. **Figure 47**  
 56. **Figure 48**  
 57. **Figure 49**  
 58. **Figure 50**  
 59. **Figure 51**  
 60. **Figure 52**  
 61. **Figure 53**  
 62. **Figure 54**  
 63. **Figure 55**  
 64. **Figure 56**  
 65. **Figure 57**  
 66. **Figure 58**  
 67. **Figure 59**  
 68. **Figure 60**  
 69. **Figure 61**  
 70. **Figure 62**  
 71. **Figure 63**  
 72. **Figure 64**  
 73. **Figure 65**  
 74. **Figure 66**  
 75. **Figure 67**  
 76. **Figure 68**  
 77. **Figure 69**  
 78. **Figure 70**  
 79. **Figure 71**  
 80. **Figure 72**  
 81. **Figure 73**  
 82. **Figure 74**  
 83. **Figure 75**  
 84. **Figure 76**  
 85. **Figure 77**  
 86. **Figure 78**  
 87. **Figure 79**  
 88. **Figure 80**  
 89. **Figure 81**  
 90. **Figure 82**  
 91. **Figure 83**  
 92. **Figure 84**  
 93. **Figure 85**  
 94. **Figure 86**  
 95. **Figure 87**  
 96. **Figure 88**  
 97. **Figure 89**  
 98. **Figure 90**  
 99. **Figure 91**  
 100. **Figure 92**  
 101. **Figure 93**  
 102. **Figure 94**  
 103. **Figure 95**  
 104. **Figure 96**  
 105. **Figure 97**  
 106. **Figure 98**  
 107. **Figure 99**  
 108. **Figure 100**  
 109. **Figure 101**  
 110. **Figure 102**  
 111. **Figure 103**  
 112. **Figure 104**  
 113. **Figure 105**  
 114. **Figure 106**  
 115. **Figure 107**  
 116. **Figure 108**  
 117. **Figure 109**  
 118. **Figure 110**  
 119. **Figure 111**  
 120. **Figure 112**  
 121. **Figure 113**  
 122. **Figure 114**  
 123. **Figure 115**  
 124. **Figure 116**  
 125. **Figure 117**  
 126. **Figure 118**  
 127. **Figure 119**  
 128. **Figure 120**  
 129. **Figure 121**  
 130. **Figure 122**  
 131. **Figure 123**  
 132. **Figure 124**  
 133. **Figure 125**  
 134. **Figure 126**  
 135. **Figure 127**  
 136. **Figure 128**  
 137. **Figure 129**  
 138. **Figure 130**  
 139. **Figure 131**  
 140. **Figure 132**  
 141. **Figure 133**  
 142. **Figure 134**  
 143. **Figure 135**  
 144. **Figure 136**  
 145. **Figure 137**  
 146. **Figure 138**  
 147. **Figure 139**  
 148. **Figure 140**  
 149. **Figure 141**  
 150. **Figure 142**  
 151. **Figure 143**  
 152. **Figure 144**  
 153. **Figure 145**  
 154. **Figure 146**  
 155. **Figure 147**  
 156. **Figure 148**  
 157. **Figure 149**  
 158. **Figure 150**  
 159. **Figure 151**  
 160. **Figure 152**  
 161. **Figure 153**  
 162. **Figure 154**  
 163. **Figure 155**  
 164. **Figure 156**  
 165. **Figure 157**  
 166. **Figure 158**  
 167. **Figure 159**  
 168. **Figure 160**  
 169. **Figure 161**  
 170. **Figure 162**  
 171. **Figure 163**  
 172. **Figure 164**  
 173. **Figure 165**  
 174. **Figure 166**  
 175. **Figure 167**  
 176. **Figure 168**  
 177. **Figure 169**  
 178. **Figure 170**  
 179. **Figure 171**  
 180. **Figure 172**  
 181. **Figure 173**  
 182. **Figure 174**  
 183. **Figure 175**  
 184. **Figure 176**  
 185. **Figure 177**  
 186. **Figure 178**  
 187. **Figure 179**  
 188. **Figure 180**  
 189. **Figure 181**  
 190. **Figure 182**  
 191. **Figure 183**  
 192. **Figure 184**  
 193. **Figure 185**  
 194. **Figure 186**  
 195. **Figure 187**  
 196. **Figure 188**  
 197. **Figure 189**  
 198. **Figure 190**  
 199. **Figure 191**  
 200. **Figure 192**  
 201. **Figure 193**  
 202. **Figure 194**  
 203. **Figure 195**  
 204. **Figure 196**  
 205. **Figure 197**  
 206. **Figure 198**  
 207. **Figure 199**  
 208. **Figure 200**  
 209. **Figure 201**  
 210. **Figure 202**  
 211. **Figure 203**  
 212. **Figure 204**  
 213. **Figure 205**  
 214. **Figure 206**  
 215. **Figure 207**  
 216. **Figure 208**  
 217. **Figure 209**

converting enzyme (ACEV), useful in identifying candidate compounds capable of binding to the variant and to detect anti-variant antibodies

Claim 1; Page 366; 519pp; English.

The sequence represents a DNA encoding an angiotensin converting enzyme splice variant (ACEV) polypeptide. The polypeptides of the invention include variants of granulocyte colony stimulating factor receptor, glucagon, interleukin 6, platelet-derived endothelial cell growth factor, cyclin-dependent kinase inhibitor 1C, cellular tumour antigen p53, and vasovactive intestinal polypeptide receptor 2. The polypeptides and their associated nucleic acids are useful for identification of variant sequences and detection of candidate compounds capable of binding the molecules. The sequences of the invention can be used in the treatment and diagnosis of various disorders including cardiovascular diseases such as arteriosclerosis, myocardial infarction and coronary arterial thrombosis, renal diseases such as diabetic nephropathy, muscular diseases such as hypertrophy, immune disorders such as immune complex nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such as asbestosis and vascular pathologies involving an endothelial abnormality such as deep vein thrombosis.

Sequence 1302 BP; 341 A; 305 C; 345 G; 311 T; 0 other;

Every Match	35.5%;	Score 250.4;	DB 22;	Length 1302;	.
Local Similarity	75.4%;	Pred. No. 6.1e-64;			
Mismatches 338;	Conservative	0;	Mismatches 106;	Indels 4;	Gaps 2;

37 CTGGGGCTACCGTCATGGATGCTCTATCAGAAGCAATGGCACATTTGCATTAAACCTT 96

60 CTTGGGTCTGCCATCATGGA TTTCTCGCAGAAGCAATGGCACCTTTGCGCTTAACCTT 11

97 TTGAAAAGCTAGGGGAAAACAACCTCAACAACCTATTTTTTC-CCCATGAGCATATCA 155

120 TTGAAAACGCTGGGTAAAGACAACTCGAAGAA TGTGTTTTTCTCACC CATGAGCATGTCC 179

156 TCAGCCCTTGGCCA TGGTTTTCATGGGGCCAAAGGGGAAACA CTCACGCTCACATCTCTCAC 315

180 TGTGGCCCTGGCCATGGTCTATGGGGCCATGGGCAACACCCCTCCACACATCCCCC 338

[illegible]

245

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

100

.....

QY 336 TTGGAGAAAAGTCTTATGATTCTCTACAGTCTTTTACAGATTCCTGTGGCAAAATCTTAC 395  
 DB 357 TTGGGAAAAGTCTGTGATTTCTCTCACTTTTAGAGATTCCTGCCAAAATCTTAC 416  
 QY 396 CAAGCAACGATAAAAACAGCTAGACTTTTGTGAATGATACAGAGAACTCCACACACGTGTA 455  
 DB 417 CAAGCAGAGATGGAGAGCTTGACTTTATCAGCGCGGTAGAGAGTCCAGAAAACACATA 476  
 QY 456 AACTCTGGTGGTGTGATAAAACTAAAG 483  
 DB 477 AACACTGGGTAGCTGAAAAGACAGAAG 504

## RESULT 12

ID AAF23495 standard; DNA; 1366 BP.

AC AAF23495;

DT 22-MAR-2001 (first entry)

XX Thrombin inhibitor DNA.

XX Thrombin inhibitor; neurodegenerative disease; Alzheimer's;

XX Parkinson's; tumour; deep venous thrombosis; pulmonary embolism;

XX heart; ss.

XX Unidentified.

OS US6156540-A.

XX 05-DEC-2000.

XX 22-DEC-1993; 93US-0171817.

XX 22-DEC-1993; 93US-0171817.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Fuldner RA, Cao L, Adams MD;

XX WPI; 2001-060165/07.

XX Novel polynucleotide encoding thrombin inhibitor polypeptide useful for

XX diagnosis and treatment of diseases related to thrombosis -

XX Claim 13; Fig.1; 19pp; English.

XX The present invention relates to a thrombin inhibitor. This is useful

XX for diagnosis and treatment of thrombosis, neurodegenerative diseases

XX such as Alzheimer's disease and Parkinson's disease, tumor metastasis,

XX to prevent proximal extension of deep venous thrombosis or the

XX recurrence of pulmonary embolisms and also to treat risk patients who

XX have congestive heart failure, acute myocardial infarction or

XX cardiomyopathy to prevent the development of deep vein thrombosis or

XX pulmonary embolism. The thrombin inhibitor is also useful for

XX screening chemical compounds that either up or down regulate its

XX activity.

XX Sequence 1366 BP; 360 A; 332 C; 355 G; 317 T; 2 other;

XX Query Match 35.3%; Score 249.2; DB 22; Length 1366;

XX Best Local Similarity 73.4%; Pred. No. 1.4e-63;

XX Matches 345; Conservative 1; Mismatches 120; Indels 4; Gaps 2;

XX 15 TCCTGGCAGCTCGCGGAGAACTGGGGCTCCCGTCATGATGCTTATCAGAGCAAA 74

XX 40 TTCTGCTCGCTCGCTCCCGCTCTGGAGTCTGCCATCATGATGTTCTGCAGAGCAAA 99

XX 75 TGGCATTGTCATTAAACCTTTTGAAGAAAGTCTAGGGGAAAACAACTCAACAACTTAT 134

XX 100 TGGCATTGCTTAAACCTTTTGAAGAAAGTCTAGGGGAAAACAACTCAACAACTTAT 159

QY 135 TTTTTC-CCCATGAGCATATCATCAGCCTTTGGCCATGGTTTTTCATGGGGCAAAAGGAAA 193  
 DB 160 TTTTCTACCCCATGAGCATGCTGTGCTTGGCCATGGTTCTACATGGGGCAAAAGGAAA 219  
 QY 194 CACTGCAGCTCAGATGCTCAGGCACCTTTTGTAGTAAATCGAGGTGAAGATGAGA 253  
 DB 220 CACGCTGCACAGATGCCAGATACCTTTCTTCAATAAAAGTGGCGGTG--GTGGAGA 276  
 QY 254 TATTTCATCGAGGTTTTTTCAGTCACCTTTTGTGCAATTAACAGAACTGACACTGAATATGT 313  
 DB 277 CATCCACAGGGCTTCCAGTCTTTTCTACCGAAGTGAACAGACTGGCAGCGAGTACTT 336  
 QY 314 GCTTAGAACTGCCAACGGGCTTTTGGAGAAAAGCTTTATGATTTCTTCACAGGTTTAC 373  
 DB 337 GCTTAGGATGCCAACAGGCTTTTGGGAAAAGCTTTGTGATTTCTCTCATCTTTAG 396  
 QY 374 AGATTCCTGTGGCAAAATTTTACCAAGCAACGATAAAACAGCTAGACTTTGTGAATGATAC 433  
 DB 397 AGATTCCTGTGGCAAAATTTTACCAAGCAAGAGATGGAGAGCTTGACTTTATCAGCGCGT 456  
 QY 434 AGAGAAGTCCCAACACAGCTGTAAACTCTCTGGTTTGTGATAAAACTAAAG 483  
 DB 457 AGAGAAGTCCCAACACACATATAACACTGGGTAGCTGAAAAGACAGAG 506

## RESULT 13

AAC65737

ID AAC65737 standard; cDNA; 1366 BP.

XX AAC65737;

XX 16-FEB-2001 (first entry)

XX Human thrombin inhibitor cDNA.

XX Thrombin inhibitor; human; thrombolytic; cardiant; cytostatic; nootropic;

XX neuroprotective; antiparkinsonian; gene therapy; treatment; carcinoma;

XX cell hypercoagulation; deep venous thrombosis; pulmonary embolism;

XX cerebral embolism; thrombic disease; tumor metastases; clot accretion;

XX neurodegenerative disease; Alzheimer's disease; Parkinson's disease; ss.

XX Homo sapiens.

XX US6133422-A.

XX 17-OCT-2000.

XX 30-NOV-1998; 98US-0200965.

XX 22-DEC-1993; 93US-0171817.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Fuldner RA, Rosen CA, Cao L, Adams MD;

XX WPI; 2001-006150/01.

XX P-PSDB; AAB11125.

XX New thrombin inhibiting proteins for treating or preventing

XX thrombosis-related diseases, e.g. embolism, tumor metastases,

XX carcinoma, neurodegenerative diseases, e.g. Alzheimer's or Parkinson's

XX disease -

XX Example 1; Column 15-16; 19pp; English.

XX This invention describes a novel human thrombin inhibitor which has

XX thrombolytic, cardiant, cytostatic, nootropic, neuroprotective and

XX antiparkinsonian activity and can be used for gene therapy. The thrombin

XX inhibitor may be used therapeutically or diagnostically to treat or

XX prevent diseases related to thrombosis characterized by hypercoagulation

XX of cells. The protein may also be used to prevent extension of deep

XX venous thrombosis of the recurrence of pulmonary embolisms and

XX recurrence of cerebral or other systemic embolisms, to treat high risk

patients, e.g. those who have congestive heart failure, acute myocardial infarction or cardiomyopathy to prevent the development of deep vein thrombosis or pulmonary embolism, and as long term therapy for occasional patient who has recurrent thrombosis or embolism while on the drug warfarin. The protein may also be used for treating thrombotic disease, tumor metastases, carcinoma of the brain, liver, lung, bone and neoplastic plasma cell carcinoma, and neurodegenerative diseases (e.g. Alzheimer's or Parkinson's disease). It can also be used against clot-bound thrombin, which produces clot accretion, for screening chemical compounds that either up or down regulate its activity, and for producing antibodies. The thrombin inhibitor can be used as pharmaceutical to mammals, such as humans, mice, rats, hamster, dogs, rabbits and cats. The cDNA sequences of the thrombin inhibitor may be used as diagnostic probe, for chromosome identification.

Sequence 1366 BP; 360 A; 332 C; 355 G; 317 T; 2 other;

Query Match 35.3%; Score 249.2; DB 22; Length 1366;  
Best Local Similarity 73.4%; Pred. No. 1.4e-63;  
Matches 345; Conservative 1; Mismatches 120; Indels 4; Gaps 2;

15 TCTGGGAGCTGCGCGAGAACTGGGGCTCAGCGTCATGATGCTCTATCAGAACAA 74  
40 TTTCTGCTCGCTGCCGCTCTGGAGTCTGCCATCATGATGTTCTCGAGAACAA 99  
75 TGGCACATTTGCATTAAACCTTTTGAAGAGCTAGGGGAAAAACAACCTCAAACTTATT 134  
100 TGGCACCTTTGGCTTTAAACCTTTTGAAGACACTGGGTAAAGACAACTCGAAGATGTGT 159  
135 TTTTTC-CCATGAGCATATCATCAGCTTTGGCCATGGTTTTCATGGGGCAAGGAAA 193  
160 TTTCTACCATCAGCATGCTCTGTGCTGGCCCTGGCCATGGTCTACATGGGGGCAAGGGAAA 219  
194 CACTGAGCTCAGATGCTCAGGCATCTTTTGTAGTAAATCGAGGTGAAGATGGAGA 253  
220 CACCGTGCACATGAGCGCCAGATCTTTCTTTCAATAAAGTGGCGTG---GTGGAGA 276  
254 TATTTCATCGAGGTTTTCAGTCACTCTCTGTCGATTAACAGACACTGACATGATGT 313  
277 CATCCACGAGGTTTCAGTCTCTCTCCAGAGTGAACAGACTGGCAGCAGTACTT 336  
314 GTTTAGAACTGCCAAGCGGCTCTTTGGAGAAAAGTCTTATGATTTTCTCAGAGTTTAC 373  
337 GCTTAGGATGGCCACAGGCTCTTTGGGAAAAGTCTTGATTTCTCTCATCTTTAG 396  
374 AGATTCCTGTGGCAATCTTACCAACAGCAATAAACAGCTAGACTTTGTGAATATAC 433  
397 AGATTCCTGCCAAAATTTTACCAACAGATGGAGGCTTGACTTTATCAGGCGGT 456  
434 AGAGAAGTCCACACAGCTGTAAACTCTCTGGGTCTCTGATAAACTAAAG 483  
457 AGAGAAGTCCAGAAAACATAAACACTGGGTGGTAGCTGAAAGACAGAG 506

RESULT 14  
AAS06063  
ID AAS06063 standard; DNA; 1324 BP.  
AC AAS06063;  
XX AAS06063;  
XX AAS06063;  
DT 12-SEP-2001 (first entry)  
DE Angiotensin converting enzyme (ACEV) splice variant DNA #63.  
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;  
KW granulocyte colony stimulating factor receptor; Glucagon; hypertrophy;  
KW platelet-derived endothelial cell growth factor; cardiovascular disease;  
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C; ds;  
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;  
KW myocardial infarction; coronary arterial thrombosis; renal disease;  
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;  
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;  
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;

vascular disorder; asbestosis.  
Homo sapiens.  
WO200136632-A2.  
25-MAY-2001.  
17-NOV-2000; 2000WO-IL00766.  
17-NOV-1999; 99IL-0132978.  
10-DEC-1999; 99IL-0133455.  
(COMP-) COMPUGEN LTD.  
Levine Z, David A, Azar I, Khosravi R, Bernstein J;  
WPI; 2001-336004/35.  
P-PSDB; AAU02963.  
Novel alternative splicing variants e.g. variant of angiotensin converting enzyme (ACEV), useful in identifying candidate compounds capable of binding to the variant and to detect anti-variant antibodies  
Claim 1; Page 360; 519pp; English.  
The sequence represents a DNA encoding an angiotensin converting enzyme splice variant (ACEV) polypeptide. The polypeptides of the invention include variants of granulocyte colony stimulating factor receptor, glucagon, interleukin 6, platelet-derived endothelial cell growth factor, cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal polypeptide receptor 2. The polypeptides and their associated nucleic acids are useful for identification of variant molecules. The sequences of the invention can be used in the treatment and diagnosis of various disorders including cardiovascular diseases such as arteriosclerosis, myocardial infarction and coronary arterial diseases such as hypertrophy, immune disorders such as immune complex nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such as asbestosis and vascular pathologies involving an endothelial abnormality such as deep vein thrombosis.  
Sequence 1324 BP; 364 A; 287 C; 360 G; 313 T; 0 other;

Query Match 35.0%; Score 246.8; DB 22; Length 1324;  
Best Local Similarity 75.8%; Pred. No. 7.2e-63;  
Matches 332; Conservative 0; Mismatches 102; Indels 4; Gaps 2;

47 CCCTCATGATGCTCTATCAGAGCAATGCGACATTTGCATTAACCTTTTGAAGAACG 106  
204 CCATCATGATGTTCTCGCAGAGCAATGCGACCTTTGCTTTAAACCTTTTGAAGAACG 263  
107 TAGGGGAAAACAACTCAAACTTATTTTTC-CCCATGAGCATATCATCAGCTTGG 165  
264 TGGCTAAAGACAACTCGAAGATGTGTTTCTCACCACATGAGCATGCTGTGCGCTGG 323  
166 CCATGGTTTTCATGGGGCAAGGAAACACACACGCTGCACAGATGGCCAGATCTTCTT 225  
324 CCATGGTCTACATGGGGCAAGGAAACACACGCTGCACAGATGGCCAGATCTTCTT 383  
226 TTAGTAAATCGAGGTGAAGATGAGATATTCATCGAGGTTTTCAGTCACTTCTTGTG 285  
384 TCAATAAAGTGGCGGT---GGTGGAGACATCCACAGGGCTTCCAGTCTTCTTCACCG 440  
286 CAATTAAAGAACTGACACTGAATATGTCTTAGAACTGCCAAGGGCTCTTTGGAGAAA 345  
441 AAGTGAACAAGACTGGCAGCAGTACTTGTCTAGGGTGGCCAAACAGAGCTCTTTGGGAAA 500  
346 AGTCTTATGATTTCTCAGAGTTTACAGATTTCTGTGGCAAAATTTTACCAAGCAACGA 405

CC- thrombosis.  
XX Sequence 1349, BP; 370 A; 298 C; 358 G; 323 T; 0 other;  
SQ Query Match 35.0%; Score 246.8; DB 22; Length 1349;  
Best Local Similarity 75.8%; Pred. No. 7.2e-63;  
Matches 332; Conservative 0; Mismatches 102; Indels 4; Gaps 2;  
QY 47 CCGTCATGATGCTCTATCAGAAAGCAAAATGGCACAATTTGCAATTAACACCTTTTGAAGAACG 106  
DB 204 CCATCATGATGTTCTCGCAGAAAGCAATGGCACCTTTTGCCTTAAACCTTTTGAAGAACG 263  
QY 107 TAGGGGAAAAACCAACTCAAACTTATTTTTC-CCATGAGCATATCATCAGCTTGG 165  
DB 264 TGGGTAAGACAACTCGAAGAAATGTGTTTTTCTCACCCTATGAGCATGCTGTGCCCTGG 323  
QY 166 CCAATGTTTTCATGGGGCAAGGGAACACACTGCAGCTCAGATGCTCAGGCACTTTGTT 225  
DB 324 CCATGCTCTACATGGGGGCAAGGGAACACCGCTGCACAGATGGCCAGATACCTTTCT 383  
QY 226 TTAGTAAAAATCGGAGGTGAAGATGGAGATATTCATCGAGGTTTTCAGTCACCTTCTGTTG 285  
DB 384 TCATTAAGAGTGGCGT---GGTGGAGACATCCACAGGGCTTCCAGTCTCTTCTCACCG 440  
QY 286 CAATTAACAGAACTGACACTGAATATGCTTAGAACTGCCAAGGGCTCTTTGGAGAAA 345  
DB 441 AAGTGAACAAAGACTGGCAGCGAGTACTTGTCTAGGGTGGCCAAACAGGCTCTTTGGGAAA 500  
QY 346 AGTCTTATGATTTTCTCACAAGGTTTTCAGATTTCTGTGGCAAAATTTTACCAGCAACGA 405  
DB 501 AGTCTGTGATTTCTCTCTCATCTTTTAGAGATTTCTGCCAAAATTTTACCAGCAGAGA 560  
QY 406 TAAACACAGCTAGACTTTTGTGAATGATACAGAGAAAGTCCACACACGTGTAACTCCTGGG 465  
DB 561 TGGAGGAGCTTGACTTTATCAGCGCGCTAGAGAAAGTCCAGAAAACACATAAACACCTGGG 620  
QY 466 TTGCTGATAAACTAAAG 483  
DB 621 TAGCTGAAAAGACAGAAG 638

Search completed: May 21, 2003, 17:38:01  
Job time : 263 secs

Db 501 AGTCTGTGATTTCTCTCATCTTTTAGAGATTTCTGCCAAAATTTTACCAGCAGAGA 560  
QY 406 TAAACACAGCTAGACTTTTGTGAATGATACAGAGAAAGTCCACACACGTGTAACTCCTGGG 465  
DB 561 TGGAGGAGCTTGACTTTATCAGCGCGCTAGAGAAAGTCCAGAAAACACATAAACACCTGGG 620  
QY 466 TTGCTGATAAACTAAAG 483  
DB 621 TAGCTGAAAAGACAGAAG 638

RESULT 15  
AAS06022  
ID AAS06022 standard; DNA; 1349 BP.  
XX  
AC AAS06022;  
XX  
XX 12-SEP-2001 (first entry)  
XX  
DE Angiotensin converting enzyme (ACEV) splice variant DNA #22.  
XX  
XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;  
XX granulocyte colony stimulating factor receptor; glucagon; hypertrophy;  
XX platelet-derived endothelial cell growth factor; cardiovascular disease;  
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C; ds;  
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;  
KW myocardial infarction; coronary arterial thrombosis; renal disease;  
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;  
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;  
KW nonrheumatic pulmonary granulomatous disease; endothelial abnormality;  
KW vascular disorder; asbestosis.

XX Homo sapiens.  
XX WO200136632-A2.  
XX  
XX 25-MAY-2001.  
XX  
XX 17-NOV-2000; 2000WO-IL00766.  
XX  
XX 17-NOV-1999; 99IL-0132978.  
PR 10-DEC-1999; 99IL-0133455.  
XX  
XX (COMP-) COMPUEN LTD.  
XX  
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;  
XX WPI; 2001-336004/35.  
XX P-PSDB; AAU02922.  
XX  
XX Novel alternative splicing variants e.g. variant of angiotensin  
XX converting enzyme (ACEV), useful in identifying candidate compounds  
XX capable of binding to the variant and to detect anti-variant antibodies

XX  
XX Claim 1; Page 327; 519pp; English.  
XX  
XX The sequence represents a DNA encoding an angiotensin converting enzyme  
XX splice variant (ACEV) polypeptide. The polypeptides of the invention  
XX include variants of granulocyte colony stimulating factor receptor,  
XX glucagon, interleukin 6, platelet-derived endothelial cell growth factor,  
XX cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P53, and  
XX vasoactive intestinal polypeptide receptor 2. The polypeptides and their  
XX associated nucleic acids are useful for identification of variant  
XX sequences and detection of candidate compounds capable of binding the  
XX molecules. The sequences of the invention can be used in the treatment  
XX and diagnosis of various disorders including cardiovascular diseases such  
XX as arteriosclerosis, myocardial infarction and coronary arterial  
XX thrombosis, renal diseases such as diabetic nephropathy, muscular  
XX diseases such as hypertrophy, immune disorders such as immune complex  
XX nephritis, multiple sclerosis, cancer, sarcoidosis, nonrheumatic  
XX pulmonary granulomatous diseases such as asbestosis and vascular  
XX pathologies involving an endothelial abnormality such as deep vein

ult No.	Score	Query Match	Length	DB	ID	Description
1	249.2	35.3	1366	3	US-09-200-965-1	Sequence 1, Appli
2	214.4	30.4	1425	1	US-08-464-148-1	Sequence 1, Appli
3	214.4	30.4	1425	1	US-08-385-500-1	Sequence 1, Appli
4	214.4	30.4	1425	1	US-08-846-784-1	Sequence 1, Appli
5	177.4	25.1	1393	1	US-08-464-148-3	Sequence 3, Appli
6	177.4	25.1	1393	1	US-08-385-500-3	Sequence 3, Appli
7	177.4	25.1	1393	1	US-08-846-784-3	Sequence 3, Appli
8	103.2	14.6	1482	1	US-07-768-286B-5	Sequence 5, Appli
9	84.8	12.0	1245	4	US-09-266-910-1	Sequence 1, Appli
10	84.8	12.0	1245	4	US-09-266-910-2	Sequence 2, Appli
11	84.8	12.0	1711	1	US-08-568-147B-1	Sequence 1, Appli
12	77.4	11.0	1316	1	US-07-753-461A-12	Sequence 12, Appli
13	77.4	11.0	1316	1	US-08-313-831A-12	Sequence 12, Appli
14	77.4	11.0	1316	1	US-08-663-318-12	Sequence 12, Appli
15	77.4	11.0	1316	5	PCF-US95-12509-12	Sequence 12, Appli
16	76.2	10.8	1152	1	US-08-315-831A-13	Sequence 13, Appli
17	76.2	10.8	1152	1	US-08-663-318-13	Sequence 13, Appli
18	76.2	10.8	1152	5	PCF-US95-12509-13	Sequence 13, Appli
19	70	9.9	1950	2	US-08-472-659-30	Sequence 30, Appli
20	70	9.9	1950	2	US-08-474-661-30	Sequence 30, Appli
21	70	9.9	1950	2	US-08-611-977-30	Sequence 30, Appli
22	65.4	9.3	1308	4	US-09-123-912-109	Sequence 109, App
23	65.4	9.3	1308	4	US-09-643-597-109	Sequence 109, App
24	65.4	9.3	1419	4	US-09-123-912-111	Sequence 111, App
25	65.4	9.3	1419	4	US-09-643-597-111	Sequence 111, App
26	63.2	9.0	1512	1	US-07-768-286B-3	Sequence 3, Appli
27	62.4	8.8	1328	1	US-07-679-052A-14	Sequence 14, Appli

QY 15 TCCTGGGCAGCTGCGCGGAGAACTGGGGCTCACCGTCATGGATGCTCTATCAGAGCAGA 74

40	TTTCTGCTCGCTCGCTCCCCGCTCTGGAGTCTGCCATCATGGATGTCTTCGCAAGACGAAA	99
75	TGGCACATTTGCATTTAAACCTTTTGA AAAAGCTAGGGGAAAA CAAC TCAAA CAAC TTA TT	134
100	TGGCACCTTTGGCTTTAAACCTTTTGA AAA CACTGGGTAAAGACAAC TCTCGAAGATGTGTT	159
135	TTTTTC - CCCATGAGCATATCATCAGCGCTTGGCCATGTGTTTTCATCTGGGGCCAAAGGGAAA	193
160	TTTTCTACCCCATGAGCATGTCTCTGTGCCCTGGCCATGGTCTCATATGGGGCCAAAGGGAAA	219
194	CACTCGAGCTCAGATGCTCAGGCACACTTTGTTTGTAGTAAAAATCGGAGGTGAAGATGGAGA	253
220	CACCGCTGCAGAGATGGGCCAGATAC TTTCTTTCAATAAAGTGGCGGTG -- GTGGAGA	276
254	TATTCATCGAGGTTTTCAGTCACTCTTTGTTGTGCAATTAACAGAACTGCACATGGAATGTT	313
277	CATCACACCGGGCTTCCAGTCTCTCTCACCGAAGTGAA CAAGACTGGCACGCAGTACTT	336
314	GCTTAGAACTGCGAAGCGGCTCTTTGGAGAAAAGTCTTATGATTTCTCTCACAGGTTTTAC	373
337	GCTTAGGATGGCCAA CAGGCTCTTTGGGGAAAAGCTTTGTGATTTCTCTCATCTTTTAG	396
374	AGATCTCTGTGGCAAATTTCAACGAAGCAACGATAAAA CAGCTAGACTTTCTGTAATGATAC	433
397	AGATTTCTGCCAAAATTTCTACCAAB CAGAGATGGAGGAGCTTGACTTTATCAGCGCGGT	456
434	AGAGAAGTCCACAA CAGCTGTAAACTCTCTGGGTTGCTGTATAAACA TAAAG	483
457	AGAGAAGTCCAGAAACACATAAACACACTGGGTAGCTGAA AAGACAGAA	506

RESULT 2  
US-08-464-148-1  
Sequence 1, Application US/08464148  
Patent No. 5710026  
GENERAL INFORMATION:  
APPLICANT: Sprecher, Cindy A.  
TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND  
TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Steuart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,148  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/385,500  
FILING DATE: 08-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 13952-21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1425 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA

:	FEATURE:				
:	NAME/KEY:	CDS			
:	LOCATION:	92..1213			
:	OTHER INFORMATION:	/product= "CYTOPLASMIC			
:	OTHER INFORMATION:	ANTIPROTEINASE-2 PROTEIN"			
:	US-08-464-148-1				
<hr/>					
	Query Match	30.4%;	Score 214.4;	DB 1;	Length 1425;
	Best Local Similarity	73.3%;	Pred. No. 3.8e-55;		
	Matches 321;	Conservative 0;	Mismatches 101;	Indels 16;	Gaps 3;
<hr/>					
QY	50	TCATGGATGCTCTATCAGAACGAATAATGGCACATTGTGCATTTAAACCTTTTGGAAAAAGCTAG	109		
Db	90	TGATGGATGACCTCTGTGTAAGCAAAATGGCACTTTTGCCATCAGCTATTATTAAAATATTGG	149		
QY	110	GGGAAA---ACNACTCAAAACAATTAATTTTTTC--CCATGAGCATATCATCAGCCTTGG	165		
Db	150	GGGAAGAGGACAACTCAAGAACAATTAATCTTCTCTCCATGAGCATCTCCTCGCCCTGG	209		
QY	166	CCATGGTTTTTCATGGGGGCAAAAGGGAAAAACACTGCAGCTCAGATGTCTCAGGCACCTTTGTT	225		
Db	210	CCATGGTCTTCATGGGGGCAAAAGGAAGCACTGCAGCCCAGATGTCCCAGGCACCTTTGTT	269		
QY	226	TTAGTAAATTCGGAGGTGAAGATGGAGATATTATCATCGAGGTTTTTCAGTCACCTTCTTGTG	285		
Db	270	T-----ATACAAAGACGGAGATATTCACCGAGTTTTCCAGTCACCTTCTCAGTG	317		
QY	286	CAATTAAACAGAACTGACACTGCACTGCAATATGTGCTTAGAATCTCCAACGGGCTTTTGAGAAA	345		
Db	318	AAGTTAAACAGAACTGGCACTCAGTACTTGTCTAGAACTGCCAACAGACTCTTTTGAGAAA	377		
QY	346	AGTCTTATGATTTCTCTACAGGTTTTTACAGATTCCTGTGGCAAATTTCTACCAAGCAACGA	405		
Db	378	AGACGTGTGATTTCTCTCCAGACTTTAAAGAAAT'ACTGT'CAGAAGTTTCTATCAGGCAGAC	437		
QY	406	TAAACACCTAGACTTTTGTGTAATGATACAGAGAAGTCCCAACACAGCTGT'ATAACTCCTGGG	465		
Db	438	TGGAGAGTTGTCTTTTGTCTCAAGACACTCAAGAGTGCAGGAAGCATATAAATGACTGGG	497		
QY	466	TTGCTGATAAAACTAAAG	483		
Db	498	TGGCAGAGAAGACTGAAG	515		

RESULT 3  
US-08-385-500-1  
Sequence 1, Application US/08385500  
Patent No. 5712117  
GENERAL INFORMATION:  
APPLICANT: Sprecher, Cindy A.  
TITLE OF INVENTION: CYTOPLASMIC ANTIPTROTEINASE-2 AND  
TITLE OF INVENTION: CYTOPLASMIC ANTIPTROTEINASE-3 AND CODING SEQUENCES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Steuart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/385.500  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 13952-21



TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1425 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 92..1213  
OTHER INFORMATION: /product= "CYTOPLASMIC  
OTHER INFORMATION: ANTIPROTEINASE-2 PROTEIN"  
US-08-395-500-1

Query Match 30.4%; Score 214.4; DB 1; Length 1425;  
Local Similarity 73.3%; Pred. No. 3.8e-55;  
Matches 321; Conservative 0; Mismatches 101; Indels 16; Gaps 3;

QY	50	TCATGGATGCTCTATCAGAACCAATGGCACTTTGGCACTTTTGAATAAGCTAG	109
DB	90	TGATGGATGACCTCTGTGAAGCAATGGCACTTTTGGCATCAGCTTATTAAATATTGG	149
QY	110	GGGAAA---ACAACCTCAAAACAACCTATTTTTTTC-CCCATGAGCATATCATCAGCCTTGG	165
DB	150	GGGAAGAGGACNACTCAAGAAACGTATTTCTCTCCATGAGCATCTCTCTGCCCTGG	209
QY	166	CCATGGTTTTCATGGGGGCAAAAGGAAACACTGCAGCTCAGATGTCTCAGGCACCTTTGTT	225
DB	210	CCATGGTCTTTCATGGGGGCAAAAGGAAAGCACTGCAGCCAGATGTCCAGGCACCTTTGTT	269
QY	226	TTAGTAAATCGGAGGTGAAGATGGAGATATTCATCGAGTTTTCAGTCACTTCTCGTTG	285
DB	270	T-----ATCAAAGAGCGAGATATTCACCGAGTTTCCAGTCACCTTCTCAGTG	317
QY	286	CAATTAACAGAACTGCACACTGAATATATGTGTTAGAACTGCCAAGGCTCTTTGGAGAAA	345
DB	318	AAGTTAAAGAACTGGCACTCAGTACTTGCTTAGAACTGCCAAGAGACTCTTTGGAGAAA	377
QY	346	AGTCTTATGATTTTCTCAGAGGTTTTCACAGATTCCTGTGGCAATTCACCAAGCAACGA	405
DB	378	AGACGTGTGATTTTCTTCCAGACTTTAAAGAACTGTCAAGAGTTCTATCAGGCAGAGC	437
QY	406	TAAACAGCTAGACTTTGTGAATGATACAGAGAGTCCCAACACGCTGTAACTCTCTGGG	465
DB	438	TGGAGGAGTTGTCTTTGCTTGAAGACACTGAAGAGTGCAGGAAGCATATAAATGACTGGG	497
QY	466	TTGCTGATAAACTAAAG	483
DB	498	TGCAGAGAAGACTGAAG	515

```

RESULT 4
US-08-846-784-1
; Sequence 1, Application US/08846784
; Patent No. 5747645
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourile and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

1 OPERATING SYSTEM:  PC-DOS/MS-DOS
2 SOFTWARE:  Patent In Release #1.0, Version #1.25
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER:  US/08/846,784
5 FILING DATE:  30-APR-1997
6 CLASSIFICATION:  435
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER:  08/385,500
9 FILING DATE:  08-FEB-1995
10 ATTORNEY/AGENT INFORMATION:
11 NAME:  Parmelee, Steven W.
12 REGISTRATION NUMBER:  31,990
13 REFERENCE/DOCKET NUMBER:  13952-21
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE:  (206) 467-9600
16 TELEFAX:  (415) 543-5043
17 INFORMATION FOR SEQ ID NO:  1:
18     SEQUENCE CHARACTERISTICS:
19     LENGTH:  1425 base pairs
20     TYPE:  nucleic acid
21     STRANDEDNESS:  single
22     TOPOLOGY:  linear
23     MOLECULE TYPE:  CDNA
24     FEATURE:
25     NAME/KEY:  CDS
26     LOCATION:  92..1213
27     OTHER INFORMATION:  /product= "CYTOPLASMIC
28     OTHER INFORMATION:  ANTIPROTEINASE-2 PROTEIN"
29 US-08-846-784-1

```

Query Match	30.4%;	Score 214.4;	DB 1;	Length 1425;
Best Local Similarity	73.3%;	Pred. No. 3.8e-55;		
Matches 321; Conservative	0;	Mismatches 101;	Indels 16;	Gaps 3;
<hr/>				
Qy	50	TCATGGATGCTCTATCAGAACGAAATGGCACATTTGCATTAAACCTTTTGGAAAAAGCTAG	109	
Db	90	TGATGGATGACCTCTGTGAAGCAATGGCACTTTTGCATCAGCTTATTATAAATATGG	149	
<hr/>				
Qy	110	GGGAAA---ACAACCTCAACCACTTATTTTTTTC-CCCATGAGCATATCATCAGCCTTGG	165	
Db	150	GGGAAGAGGACAACTCAAGAAAGATTTCTCTCTCCATGAGCATCTCCTCTGCCCTGG	209	
<hr/>				
Qy	166	CCATGGTTTTTCATGGGGGCCAAAGGGAAACACTGCAGCTCAGATGTCTCAGGCACCTTTGTT	225	
Db	210	CCATGGTCTTCATGGGGGCCAAAGGGAAGCACTGCAGCCAGATGTCCAGGGCACTTTGTT	269	
<hr/>				
Qy	226	TTAGTAAATCGGAGGTGAAGATGGAGATATTCAATCGAGGTTTTTCAGTCACTTCTTGTGG	285	
Db	270	T-----ATACAAGACGGAGATATTCAACGAGGTTTCCAGTCACTTCTCACTG	317	
<hr/>				
Qy	286	CAATTAAACGAACCTGACACTGAATATGCTTTAGAACTGCCACGGGCTCTTTGGAGAAA	345	
Db	318	AGGTTAAAGAACTGGCACTCAGTACTTCTTTAGAACTGCCAAGCACTCTTTGGAGAAA	377	
<hr/>				
Qy	346	AGTCTTATGATTTCTCAAGGTTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCAACGA	405	
Db	378	AGACGTGTGATTTCTCTCCAGACTTTTAAAGATCTGTGCAGAAATTTCTATCAGCAGAGC	437	
<hr/>				
Qy	406	TAAACACAGCTAGACTTTTGTAATGATACAGAGAGTCCACACACGTGTAAACTCTGGG	465	
Db	438	TGGAGGAGTTGTCTTTTGTCTGAAGACACTGAAAGGTGCAGGAGCATATAAAATGACTGGG	497	
<hr/>				
Qy	466	TTGCTGATAAAACTAAAG	483	
Db	498	TGGCAGAGAAGACTGAAG	515	

RESULT 5  
US-08-464-148-3  
; Sequence 3, Application US/08464148.  
; Patent No. 5710026  
; GENERAL INFORMATION:  
; APPLICANT: Sprecher, Cindy A.

Db 449 ATGCTGAGCTGAAGGAGCTTTCTTTATCAGAGCTGCAGAGAGTCCAGGAAACACATCA 508

Qy 457 ACTCCTGGTGTCTGATATAAAG 483

Db 509 ACACCTGGGTCTCAAAAAGACCCGAAG 535

RESULT 6

US-08-385-500-3

; Sequence 3, Application US/08385500

; Patent No. 5712117

GENERAL INFORMATION:

APPLICANT: Sprecher, Cindy A.

TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND

TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourie and Crew

STREET: Stuart Street Tower, One Market Plaza

CITY: San Francisco

STATE: California

COUNTRY: US

ZIP: 94105-1493

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/385,500

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 13952-21

TELEPHONE: (206) 467-9600

TELEFAX: (415) 543-5043

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1393 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 112..1239

OTHER INFORMATION: /product= "CYTOPLASMIC

OTHER INFORMATION: ANTIPROTEINASE-3 PROTEIN"

US-08-385-500-3

Query Match 25.1%; Score 177.4; DB 1; Length 1393;

Best Local Similarity 66.0%; Pred. No. 5.8e-44;

Matches 295; Conservative 0; Mismatches 136; Indels 16; Gaps 2;

Qy 41 GGCCTCACCGTCATGGATGCTCTATCAGAGCAAAATGGACATTTGCAATTAACCTTTTGA 100

Db 101 GGCCTGCATCATGGAACTCTTTCTAATGCAAGTGTACTTTTGGCATAACGCTTTTAA 160

Qy 101 AAAAGCTAGGGGAAACAACTCAACAC-----TTATTTTTCCTCCATGAGCATATCAT 156

Db 161 AGATACTGTGTCAAGATAACCCCTTCGCACAACGTTCTCTCTCTGTGAGCATCTCT 220

Qy 157 CAGCCTTGGCCATGTTTTCATGGGGCAAGGAAACACCTGCAGCTCAGATGTCACAG 216

Db 221 CTGCCCTGGCCATGTTCTCTAGGGCAAGGAAACACCGCAACCCAGATGGCCAGG 280

Qy 217 CACTTTGTTTGTAGTAAATCGGAGGTGAAGATGAGATATTCATCGAGTTTTCAGTCAC 276

Db 281 CACTGTCTTTAAACA-----CAGAGGAAGACATTCATCGGCTTTTCCAGTCGC 328

;; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND

;; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES.

;; NUMBER OF SEQUENCES: 16

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Townsend and Townsend Kourie and Crew

;; STREET: Stuart Street Tower, One Market Plaza

;; CITY: San Francisco

;; STATE: California

;; COUNTRY: US

;; ZIP: 94105-1493

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent In Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/464,148

;; FILING DATE: 05-JUN-1995

;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 08/385,500

;; FILING DATE: 08-FEB-1995

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Parmelee, Steven W.

;; REGISTRATION NUMBER: 31,990

;; REFERENCE/DOCKET NUMBER: 13952-21

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (206) 467-9600

;; TELEFAX: (415) 543-5043

;; INFORMATION FOR SEQ ID NO: 3:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 1393 base pairs

;; TYPE: nucleic acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: cDNA

;; FEATURE:

;; NAME/KEY: CDS

;; LOCATION: 112..1239

;; OTHER INFORMATION: /product= "CYTOPLASMIC

;; OTHER INFORMATION: ANTIPROTEINASE-3 PROTEIN"

US-08-464-148-3

Query Match 25.1%; Score 177.4; DB 1; Length 1393;

Best Local Similarity 66.0%; Pred. No. 5.8e-44;

Matches 295; Conservative 0; Mismatches 136; Indels 16; Gaps 2;

Qy 41 GGCCTCACCGTCATGGATGCTCTATCAGAGCAAAATGGACATTTGCAATTAACCTTTTGA 100

Db 101 GGCCTGCATCATGGAACTCTTTCTAATGCAAGTGTACTTTTGGCATAACGCTTTTAA 160

Qy 101 AAAAGCTAGGGGAAACAACTCAACAC-----TTATTTTTCCTCCATGAGCATATCAT 156

Db 161 AGATACTGTGTCAAGATAACCCCTTCGCACAACGTTCTCTCTCTGTGAGCATCTCT 220

Qy 157 CAGCCTTGGCCATGTTTTCATGGGGCAAGGAAACACCTGCAGCTCAGATGTCACAG 216

Db 221 CTGCCCTGGCCATGTTCTCTAGGGCAAGGAAACACCGCAACCCAGATGGCCAGG 280

Qy 217 CACTTTGTTTGTAGTAAATCGGAGGTGAAGATGAGATATTCATCGAGTTTTCAGTCAC 276

Db 281 CACTGTCTTTAAACA-----CAGAGGAAGACATTCATCGGCTTTTCCAGTCGC 328

Qy 277 TTCTTGTGCAATTAACAGAACTCACACTGAATATGTCTTAGAATGCCAAGCGGCTCT 336

Db 329 TTCTCACTGAAGTGAACAGGCTGGCACAGTACCTGCTGAGAACGGCCACAGGCTCT 388

Qy 337 TTGAGAGAAAGCTCTTATGATTTCTCTCAGGTTTTACAGATTCCTGTGGCAATTTTACC 396

Db 389 TTGAGAGAAAGCTTTGTCAGTTCTCTCAAGTTTAAAGGAATCTGTCTTCAATTTTACC 448

Qy 397 AAGCAACGATTAACAGCTAGACTTTTGTGATGATGATGATGATGATGATGATGATGAT 456

QY 277 TTCTTGTGCAATTAACAGAACTGACACTGAATATGTCTTAGAACTGCCAACGGGCTCT 336  
 Db 329 TTCTCAGTGAAGTGAACAAAGGCTGGCACACAGTACCTGCTGAGAAACGGCCAAAGGCTCT 388  
 QY 337 TTGAGAGAAAGTCTTATGATTTCTCAGAGGTTTACAGATTCCTGTGGCAATTTCTACC 396  
 Db 389 TTGAGAGAAAGTCTTATGATTTCTCAGAGGTTTACAGATTCCTGTGGCAATTTCTACC 448  
 QY 397 AAGCAACGATAAACAGCTAGACTTTGTGAATGATACAGAGAACTCCACAAACAGCTGTAA 456  
 Db 449 ATGCTGAGCTGAAGGAGCTTTCTTTATCAGAGCTGCAGAGAGTCCAGGAACACATCA 508  
 QY 457 ACTCCTGGTGTGATAAACTAAAG 483  
 Db 509 ACACCTGGTGTCAAAAAGACCGAAG 535

RESULT 7  
 US-07-768-286B-5  
 ; Sequence 3, Application US/08846784  
 ; Patent No. 5747645  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sprecher, Cindy A.  
 ; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND  
 ; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend Kourie and Crew  
 ; STREET: Stewart Street Tower, One Market Plaza  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: US  
 ; ZIP: 94105-1493

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/846,784  
 FILING DATE: 30-APR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/385,500  
 FILING DATE: 08-FEB-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Parmelee, Steven W.  
 REGISTRATION NUMBER: 31,990  
 REFERENCE/DOCKET NUMBER: 13952-21  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 467-9600  
 TELEFAX: (415) 543-5043  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1393 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 112..1239  
 OTHER INFORMATION: /product= "CYTOPLASMIC  
 OTHER INFORMATION: ANTIPROTEINASE-3 PROTEIN"

US-08-846-784-3  
 Query Match 25.1%; Score 177.4; DB 1; Length 1393;  
 Best Local Similarity 66.0%; Pred. No. 5.8e-44;  
 Matches 295; Conservative 0; Mismatches 136; Indels 16; Gaps 2;

QY 41 GGCTCACCGTCATGATGCTCTATCAGAGCAATGGCACATTTGCAATTAACCTTTTGA 100  
 Db 101 GGCCCTGCATCATGGAACACTCTTCTAATGCAAGTGGTACTTTTGGCCATACGCCCTTTAA 160

QY 101 AAAAGCTAGGGGAAACAACTCAAAACAC-----TTATTTTTTCCCCATGAGCATATCAT 156  
 Db 161 AGATACGTGTCAAGATAACCCCTTCGCAACAGTGTCTGTCTCTCTGTGAGCATCTCCT 220  
 QY 157 CAGCCCTTGCCCATGGTTTTCATGGGGGCAAGGAAACACCTCAGCTCAGATGTCTCAGG 216  
 Db 221 CTGCCCTGGCCATGGTTCTCTTAGGGGCAAGGAAACACCCCAACCCAGATGGCCAGG 280  
 QY 217 CACTTTGTTTAAATTCGAGGTGAAGATGAGATATTTCATCGAGGTTTTTCAGTCCAC 276  
 Db 281 CACTGTCTTTAAACA-----CAGAGGAAGACATTTCATCGGCTTTTCCAGTCGC 328  
 QY 277 TTCTTGTGCAATTAACAGAACTGACACTGAATATGTCTTAGAACTGCCAACGGGCTCT 336  
 Db 329 TTCTCAGTGAAGTGAACAAAGGCTGGCACACAGTACCTGCTGAGAAACGGCCAAAG 388  
 QY 337 TTGAGAGAAAGTCTTATGATTTCTCAGAGGTTTACAGATTCCTGTGGCAATTTCTACC 396  
 Db 389 TTGAGAGAAAGTCTTATGATTTCTCAGTTCCTCTCAACGTTTAAAGAACTCTGTCTTCAATTTCTACC 448  
 QY 397 AAGCAACGATAAACAGCTAGACTTTGTGAATGATACAGAGAACTCCACAAACAGCTGTAA 456  
 Db 449 ATGCTGAGCTGAAGGAGCTTTCTTTATCAGAGCTGCAGAGAGTCCAGGAACACATCA 508  
 QY 457 ACTCCTGGTGTGATAAACTAAAG 483  
 Db 509 ACACCTGGTGTCAAAAAGACCGAAG 535

RESULT 8  
 US-07-768-286B-5  
 ; Sequence 5, Application US/07768286B  
 ; Patent No. 5444153  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GOSS, Neil H.  
 ; APPLICANT: RICHARDSON, Michael A.  
 ; TITLE OF INVENTION: VARIANTS OF PAI-2  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 1800 Diagonal Road, Suite 500  
 ; CITY: Alexandria  
 ; STATE: Virginia  
 ; COUNTRY: USA  
 ; ZIP: 22313-0299  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/768,286B  
 FILING DATE: 19911011  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/AU90/00603  
 FILING DATE: 20-DEC-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 16786/157 CHAC  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)836-9300  
 TELEFAX: (703)683-4109  
 TELEX: 899149  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1482 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: Homo sapiens  
CELL TYPE: Monocyte  
CELL LINE: U937  
IMMEDIATE SOURCE: BTA 1922  
CLONE: BTA 1922  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 22..1170  
IDENTIFICATION METHOD: experimental  
OTHER INFORMATION: /function= "Product binds to urokinase, tissue plasminogen activator"  
OTHER INFORMATION: /product= "PAI-2 variant, protease sensitive site removed"  
OTHER INFORMATION: /evidence= EXPERIMENTAL  
OTHER INFORMATION: /notes= "Codes for human plasminogen activator inhibitor type 2 protein in which amino acids 66 to 98 inclusive have been deleted."  
OTHER INFORMATION: to 98 inclusive have been deleted."  
US-07-768-286B-5

Query Match 14.6%; Score 103.2; DB 1; Length 1482;  
Best Local Similarity 53.9%; Pred. No. 1.6e-21;  
Matches 236; Conservative 0; Mismatches 198; Indels 4; Gaps 1;

QY 52 ATGGATGCTCTATCAGAGCAATGGCACAATTTGCAATTAACCTTTTGAAGAGCTAGG 111  
DB 22 ATGGAGATCTTTGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGCA 81

QY 112 GAAACCACTCAAC-----AACTTATTTTCCCATGAGATATCATCAGCTTGGCC 167  
DB 82 AAGCAAGCCCAACCAACCACTCTCTCTCCCATGAGATCTCGTCCACCATGGCC 141

QY 168 ATGGTTTTCATGGGGCAAGGAAACACTGCACTGAGTCTCAGGACCTTTGTTT 227  
DB 142 ATGGTCTACATGGCTCCAGGGGAGCAACCAAGACAGATGGCCAAAGTGTCTCAGTTT 201

QY 228 AGTAAATCGAGGTGAGTGAATGAGATATCATCGAGGTTTTCAGTCTCTTGTGCA 287  
DB 202 AATGAAGTGGAGCGCTGAGATAAATCCATTCCTTCGCTCTCTCAGTCTGCA 261

QY 288 ATTAACAGAACTGACACTGAATATGTCTTAGAACTGCCAACGGGCTTTTGGAGAAAAG 347  
DB 262 ATCAATGATCCACAGGGAATTTTACTGGAAGTGTCAATAAGCTGTTGTGAGAAG 321

QY 348 TCTTATGATTTCTCAGCTTTTACAGATCTCTGTGGCAATTTTACCAAGCAAGATA 407  
DB 322 TCTGCGAGCTTCCGGGAAGAATATATTCGACTCTGTGAGAAATATTACTCTCAGAACCC 381

QY 408 AAACAGCTAGACTTTGTGAATGATACAGAGAAGTCCACACACGCTGTAAACTCTGGTT 467  
DB 382 CAGCAGTACCTTCTAGATGTGCAAGAGCTAGAAAAGATTATCTCTGGTC 441

QY 468 GCTGATAAACTAAAGCC 485  
DB 442 AAGACTCAACCAAGGC 459

RESULT 9  
US-09-266-910-1  
; Sequence 1, Application US/09266910  
; Patent No. 6344362  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Use of a recombinant protein as receptor of a  
; NUMBER OF SEQUENCES: 8  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/266-910  
; FILING DATE:

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/266,910  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1245 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: Squamous Cell Carcinoma Antigen  
TISSUE TYPE: Hepatoma  
CELL TYPE: Hepatocyte  
CELL LINE: HepG2  
US-09-266-910-1

Query Match 12.0%; Score 84.8; DB 4; Length 1245;  
Best Local Similarity 53.4%; Pred. No. 5.3e-16;  
Matches 251; Conservative 0; Mismatches 182; Indels 37; Gaps 2;

QY 51 CATGGATGCTCTATCAGAGCAATGGCACAATTTGCAATTAACCTTTTGAAGAGCTAGG 110  
DB 7 CATGAATTCATCTAGTGAAGCCACACCAAGTTTCATGTTCCAGCTGTTCACAGATTTCAG 66

QY 111 GGAACCACTCAACCAACTTATTTTTCCTCC-ATGAGCATATCATCAGCTTGGCCAT 169  
DB 67 AAAATCAAGAGAGAACCAATCTTCTATTTCCCTATCAGATCATCATCAGATTAGGAT 126

QY 170 GGTTCATGGGGCAAGGAAACACTGCACTGAGTCTCAGGACCTTTGTTTGTAG 229  
DB 127 GGTCTCTTAGAGCCCAAGACAACTGCAACAGATTAAAGAGGTTCTTCACTTTGA 186

QY 230 TAAATCGAGGTGAAAT-----GGAGA 253  
DB 187 TCAAGTCACAGAGAACCAACAGGAAAGTGCACATATCATGTTAGTTCAGGAAA 246

QY 254 TATTCATCGAGGTTTTCAGTCACTTCTTTGTCGAATTAACAGAACTGACATGATATGT 313  
DB 247 TGTTCATCAGGTTTCAAGAGCTTCTGACTGAATTCACAAATCCACTGATGATATGA 306

QY 314 GCTTAGAACTGCCAAAGGCTCTTTGGAGAAAAGTCTTATGATTTCTCAGGTTTAC 373  
DB 307 GCTGAAGATCCCAACAGCTCTTCGGAGAAAAACGTTATCTATTTTACAGGAATATT 366

QY 374 AGATTCTCTGCGCAATTTCTACCAAGCAACGATAAAAACAGCTAGACTTTTGTGAATGATAC 433  
DB 367 AGATGCCATCAAGAAATTTTACCAAGCCAGTGTGGAATCTGTTGATTTGCAATGCTCC 426

QY 434 AGAGAAGTCCCAACACGCTGTAAACTCTGGTTGCTGTGATAAAACTAAAG 483  
DB 427 AGAAGAAAGTCGAAGAGATTAATCTCTGGTGGAAAGTCAACAGATG 476

RESULT 10  
US-09-266-910-2  
; Sequence 2, Application US/09266910  
; Patent No. 6344362  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Use of a recombinant protein as receptor of a  
; NUMBER OF SEQUENCES: 8  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/266-910  
; FILING DATE:



Wed May 28 14:44:08 2003

```

; TITLE OF INVENTION: Human Elastase Inhibitor
; NUMBER OF SEQUENCES:
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Eileen Remold-O'Donnell, Ph.D.
; STREET: 197 Clinton Street
; CITY: Brookline
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 02146
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch
; COMPUTER: IBM-compatible
; OPERATING SYSTEM: MS-DOS Version 3.3
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,461A
; FILING DATE: 19910906
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/314,383
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Gates, Edward R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0279/7012
; INFORMATION FOR SEQ ID NO: 12
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1316 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double standard
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; US-07-755-461A-12

Query Match 11.0%; Score 77.4; DB 1; Length 1316;
Best Local Similarity 51.4%; Pred. No. 9.4e-14;
Matches 240; Conservative 0; Mismatches 211; Indels 16; Gaps 2;

QY 21 GCAGCTCGCGGAGAACTGGGGCTCACCGTCATGGATGCTCTATCAGAAGCAAAATGGCAC 80
DB 18 GGAGGCTCTCGCGGCTGTCCGTTTTCACCATGGAGCAGCTGAGCTCAGCAAAACCCG 77

QY 81 ATTTGCATTAAACCTTTTGAAAAAGCTAGGGGAAAAACAATCTC-----AAACAATTTATTTT 136
DB 78 CTTCGCTTGGACCTGTTCTCGGCTTGAGTGAGAAACAATCCGGCTGGAACATCTTCAT 137

QY 137 TTTCCCATGAGCATATCATCAGCCCTTGGCCATGTTTTCATGGGGCAAGGGAACAC 196
DB 138 CTCTCCCTTCAAGTTTCTGCTATGCGGCTGTCTGTTTCTGGGGACAGAGGTAACAC 197

QY 197 TCGAGCTCAGATGCTCTCAGGCACTTTGTTTGTAGTAAATCGAGGTGAAGATGGAGATAT 256
DB 198 GGCAGCAGCTGTCCAGACTTTCATTTCCATTTCAACAGGT-----TGAAGAGGT 245

QY 257 TCATCGAGGTTTTCAGTCACTTTCTTTGTTGCAATTAACAGAACTGAACTAAAG 483
DB 426 AGATGCAAGGAGACCATTAACCACTGGGTCAAGACACAGACAGAG 472

; TITLE OF INVENTION: Human Elastase Inhibitor
; NUMBER OF SEQUENCES:
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Eileen Remold-O'Donnell, Ph.D.
; STREET: 197 Clinton Street
; CITY: Brookline
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 02146
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch
; COMPUTER: IBM-compatible
; OPERATING SYSTEM: MS-DOS Version 3.3
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,461A
; FILING DATE: 19910906
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/314,383
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Gates, Edward R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0279/7012
; INFORMATION FOR SEQ ID NO: 12
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1316 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double standard
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; US-07-755-461A-12

Query Match 11.0%; Score 77.4; DB 1; Length 1316;
Best Local Similarity 51.4%; Pred. No. 9.4e-14;
Matches 240; Conservative 0; Mismatches 211; Indels 16; Gaps 2;

QY 21 GCAGCTCGCGGAGAACTGGGGCTCACCGTCATGGATGCTCTATCAGAAGCAAAATGGCAC 80
DB 18 GGAGGCTCTCGCGGCTGTCCGTTTTCACCATGGAGCAGCTGAGCTCAGCAAAACCCG 77

QY 81 ATTTGCATTAAACCTTTTGAAAAAGCTAGGGGAAAAACAATCTC-----AAACAATTTATTTT 136
DB 78 CTTCGCTTGGACCTGTTCTCGGCTTGAGTGAGAAACAATCCGGCTGGAACATCTTCAT 137

QY 137 TTTCCCATGAGCATATCATCAGCCCTTGGCCATGTTTTCATGGGGCAAGGGAACAC 196
DB 138 CTCTCCCTTCAAGTTTCTGCTATGCGGCTGTCTGTTTCTGGGGACAGAGGTAACAC 197

QY 197 TCGAGCTCAGATGCTCTCAGGCACTTTGTTTGTAGTAAATCGAGGTGAAGATGGAGATAT 256
DB 198 GGCAGCAGCTGTCCAAGACTTTCATTTCAACAGGT-----TGAAGAGGT 245

QY 257 TCATCGAGGTTTTCAGTCACTTTCTTTGTTGCAATTAACAGAACTGAACTGAAATGTGCT 316
DB 246 TCATTCAGATTCAGAGTCTGATGCTGATATCAACAAACGTTGGAGGCTTATATCT 305

QY 317 TAGAACTGCCAACGGGCTTTTGGAGAAAAGCTTATGATTTCTCAGAGTTTACAGA 376
DB 306 GAAACTTGTCTAATAGATTATATGAGAGAAAACCTTCAATTTCCCTCCTGAGTTCTTGGT 365

QY 377 TTCCTGTGGCAAACTTCAACAGCAACGATAAACAACAGCTAGACTTTGTGTAATGATACAGA 436
DB 366 TTCGACTCAGAAAAACATATGTTGCTGACCTGGCCAGTGTGATTTTCAGCATGCTCTGA 425

QY 437 GAAGTCCACAACACGCTGTAACCTCCTGGGTTGCTGTGATAAACTAAAG 483
DB 426 AGATGCAAGGAGACCATTAACCACTGGGTCAAGACACAGACAGAG 472

RESULT 13
US-08-315-831A-12
; Sequence 12: Application US/08315831A
; Patent No. 5663299
; GENERAL INFORMATION:
; APPLICANT: Remold-O'Donnell, Eileen
; TITLE OF INVENTION: Human Monocyte Elastase Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Ave.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,831A
; FILING DATE: September 30, 1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Plummer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: C0279/7016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1316 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 49..1188
; OTHER INFORMATION: /codon_start= 49
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 49..1185
; US-08-315-831A-12

Query Match 11.0%; Score 77.4; DB 1; Length 1316;
Best Local Similarity 51.4%; Pred. No. 9.4e-14;
Matches 240; Conservative 0; Mismatches 211; Indels 16; Gaps 2;

QY 21 GCAGCTCGCGGAGAACTGGGGCTCACCGTCATGGATGCTCTATCAGAAGCAAAATGGCAC 80
DB 18 GGAGGCTCTCGCGGCTGTCCGTTTTCACCATGGAGCAGCTGAGCTCAGCAAAACCCG 77

QY 81 ATTTGCATTAAACCTTTTGAAAAAGCTAGGGGAAAAACAATCTC-----AAACAATTTATTTT 136
DB 78 CTTCGCTTGGACCTGTTCTCGGCTTGAGTGAGAAACAATCCGGCTGGAACATCTTCAT 137

QY 137 TTTCCCATGAGCATATCATCAGCCCTTGGCCATGTTTTCATGGGGCAAGGGAACAC 196
DB 138 CTCTCCCTTCAAGTTTCTGCTATGCGGCTGTCTGTTTCTGGGGACAGAGGTAACAC 197

QY 197 TCGAGCTCAGATGCTCTCAGGCACTTTGTTTGTAGTAAATCGAGGTGAAGATGGAGATAT 256
DB 198 GGCAGCAGCTGTCCAGACTTTCATTTCCATTTCAACAGGT-----TGAAGAGGT 245

QY 257 TCATCGAGGTTTTCAGTCACTTTCTTTGTTGCAATTAACAGAACTGAACTGAAATGTGCT 316
DB 246 TCATTCAGATTCAGAGTCTGATGCTGATATCAACAAACGTTGGAGGCTTATATCT 305

QY 317 TAGAACTGCCAACGGGCTTTTGGAGAAAAGCTTATGATTTCTCAGAGTTTACAGA 376
DB 306 GAAACTTGTCTAATAGATTATATGAGAGAAAACCTTCAATTTCCCTCCTGAGTTCTTGGT 365

QY 377 TTCCTGTGGCAAACTTCAACAGCAACGATAAACAACAGCTAGACTTTGTGTAATGATACAGA 436
DB 366 TTCGACTCAGAAAAACATATGTTGCTGACCTGGCCAGTGTGATTTTCAGCATGCTCTGA 425

QY 437 GAAGTCCACAACACGCTGTAACCTCCTGGGTTGCTGTGATAAACTAAAG 483
DB 426 AGATGCAAGGAGACCATTAACCACTGGGTCAAGACACAGACAGAG 472
```

Db 246 TCATTCAAGATTCAGAGTCTGAATGCTGATATCAACAAACGCTGGAGCGTCTTATATCT 305  
 QY 317 TAGAAGTCCACCGGCTCTTTGGAGAAAGTCTTATGATTTCCCTCACAGGTTTACAGA 376  
 Db 306 GAACTTGTCTAATAGATATATGGAGAAACTTACAAATTCCTTCTGAGTCTTGCT 365  
 QY 377 TTCCTGTGGCAATTTTACCAGCAACGATATAACAGCTAGACTTTGTGAATGATACAGA 436  
 Db 366 TTCGACTCAGAAACATATGGTGTGACCTGGCCAGTGTGGATTTTCAGCATGCTCTGA 425  
 QY 437 GAAGTCCACCAACACGCTGTAACCTCTGGTGTCTGATATAAACTAAAG 483  
 Db 426 AGATCAAGGAAGACCATTAACCACTGGGTCAAGGACAGACAGAAG 472

## RESULT 14

US-08-662-318-12  
 Sequence 12, Application US/08662318  
 Patent No. 5827672

## GENERAL INFORMATION:

APPLICANT: Remold-O'Donnell, Eileen  
 TITLE OF INVENTION: Human Monocyte Elastase Inhibitor  
 NUMBER OF SEQUENCES: 16  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 STREET: 600 Atlantic Ave.  
 CITY: Boston  
 STATE: MA

## COUNTRY: USA

ZIP: 02210

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/662,318  
 FILING DATE:

## CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/315,831  
 FILING DATE: September 30, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plumer, Elizabeth R.  
 REGISTRATION NUMBER: 36,637  
 REFERENCE/DOCKET NUMBER: C0279/7016  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1316 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 49..1188

OTHER INFORMATION: /codon\_start= 49

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 49..1185

US-08-662-318-12

## Query Match

Best Match 11.0%; Score 77.4; DB 1; Length 1316;

Best Local Similarity 51.4%; Pred. No. 9.4e-14;

Matches 240; Conservative 0; Mismatches 211; Indels 16; Gaps 2;

QY 21 GCAGCTGCGGAGAACTGGGGCTCACCGTCTATCATGAGCAAAATGGCAC 80  
 Db 18 GGAAGGTCTCGCGGCTGTCGGTTCACCATGGAGCAGCTGAGCTCAGCAAAACCCCG 77  
 QY 81 ATTTCATTAAACCTTTTGAAAAAGCTAGGGGAAAAACAATC-----AAACAACCTTATTTT 136  
 Db 78 CTTTCGCTTGGACCTGTTCTCGGCGTTGAGTGAGAACAAATCCGGCTGGAAACATCTTCAT 137  
 QY 137 TTTTCCCATGAGCATATCATCAGCTTGGCCATGGTTTTCATGGGGGCAAGGAAACAC 196  
 Db 138 CTCCTCCCTTCAGCAATTCATCTGCTATGGCCATGGTTTTCCTGGGACCAGAGTAACAC 197  
 QY 197 TGCAGCTCAGATGCTCTCAGGCACTTTGTTTGTAGTAAATCGGAGGTGAAGATGGAGATAT 256  
 Db 198 GGCAGCACAGCTGTCCAAGACTTTCCATTTCAACCGGT-----TGAAGAGGT 245  
 QY 257 TCATCGAGGTTTTCAGTCACTTCTTGTTCGAATTAACAGAACTGACACTGAAATATGTGCT 316  
 Db 246 TCATTCAAGATTCAGAGTCTGAATGCTGATATCAACAAACGCGGAGCGTCTTATATCT 305  
 QY 317 TAGAAGTCCCAACGGGCTCTTTGGAGAAAGTCTTATGATTTCCCTCACAGGTTTACAGA 376  
 Db 306 GAAACTTGTCTAATAGATATATGGAGAGAAAACTTACAAATTTCTCTGAGTCTTGCT 365  
 QY 377 TTCCTGTGGCAAAATTCACCAAGCAACGATATAACAGCTAGACTTTGTGAATGATACAGA 436  
 Db 366 TTCGACTCAGAAACATATGTTGCTGACCTGGCCAGTGTGGATTTTCAGCATGCTCTGA 425  
 QY 437 GAAGTCCACACACGCTGTAACTCTGGTGTGCTGATATAAACTAAAG 483  
 Db 426 AGATCAAGGAAGACCATTAACCACTGGGTCAAGGACAGACAGAAG 472

## RESULT 15

PCT-US95-12509-12  
 Sequence 12, Application PC/TUS9512509  
 GENERAL INFORMATION:

## APPLICANT:

TITLE OF INVENTION: Human Monocyte Elastase Inhibitor  
 NUMBER OF SEQUENCES: 16  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 STREET: 600 Atlantic Ave.  
 CITY: Boston  
 STATE: MA

## COUNTRY: USA

ZIP: 02210

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/12509  
 FILING DATE: herewith

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/315,831  
 FILING DATE: 30 September 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plumer, Elizabeth R.  
 REGISTRATION NUMBER: 36,637  
 REFERENCE/DOCKET NUMBER: C0279/7016  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1316 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

Wed May 28 14:44:08 2003

```

; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 49..1188
; OTHER INFORMATION: /codon_start= 49
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 49..1185
; PCT-US95-12503-12

Query Match      11.0%; Score 77.4; DB 5; Length 1316;
Best Local Similarity 51.4%; Pred. No. 9.4e-14;
Matches 240; Conservative 0; Mismatches 211; Indels 16; Gaps 2;

Qy 21 GCAGCTGCGGGAGAACTCGGGGCTCACCGTCATGGATGCTCTATCAGAGCAAAATGGCAC 80
Db 18 GGAGCGTCTCGGGGCTGTCGGTTTTCACCATGGAGCAGCTGAGCTCAGCAAAACACCGG 77

Qy 81 ATTGCAATTAACCTTTTGAAGAAAGCTAGGGGAAACAATCTC-----AAACAATTTATTT 136
Db 78 CTTCGCCCTTGACCTGTTCTCGGGGTTGAGTGAGAAACAATCCGGCTGGAAACAATCTTCAT 137

Qy 137 TTTCCCATGAGCATATCATCAGCCTTGGCCATGGTTTTCATGGGGGCAAGGGAACAC 196
Db 138 CTCTCCCTTACGCAATTCATCTGCTATGCCATGGTTTTCGGGACCAAGAGGTAACAC 197

Qy 197 TGCAGCTCAGATGCTTCAGGCACCTTTGTTTAGTAAATCGAGGTCGAGATGAGATAT 256
Db 198 GGCAGCACAGCTGTCCAAGACTTTCACATTTCAACACGGT-----TGAAGAGGT 245

Qy 257 TCATCGAGGTTTTCAGTCACCTTCTGTGCAATTAACAGAACTCAGACCTGAATATGTCT 316
Db 246 TCATTCAAGATTCAGAGTCTGAATGCTGATATCAACAACGGTGGAGCGCTTATATCT 305

Qy 317 TAGAACTGCCAAACGGGCTCTTTGGAGAAAAGTCTTATGATTTCCCTCAGGTTTACAGA 376
Db 306 GAAACTTGCTAATAGATTATATGAGAGAAAACCTTACAAATTCCTCTGAGTCTTGCT 365

Qy 377 TTCCTGTGGCAAAATTCACCAAGCAACGATAAACAGCTAGACTTTGTGAATGATACAGA 436
Db 366 TTCGACTCAGAAAACATATGTTGCTGACCTGGCCAGTGTGGATTTTCAGCATGCCCTCTGA 425

Qy 437 GAAGTCCACACACGCTGTAACTCTCGGTTGCTGATAAAACTAAAG 483
Db 426 AGATCAAGGAAGACCATTAACCCAGTGGGTCAAAGGACAGACAGAAG 472
```

Search completed: May 21, 2003, 19:15:18  
Time : 57 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 20:05:44 ; Search time 113 Seconds  
(without alignments)  
8249.968 Million cell updates/sec

Title: US-09-912-628-4  
Perfect score: 706  
Sequence: 1 ggcacgagcttcgctctgg.....gtaaaaaaaaaaaaaa 706

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

hed: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- Published Applications NA:\*
- 1: /cgn2\_5/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/pubpna/PCTUS\_NEW\_PUB.seq:\*
  - 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	704	99.7	706	9	US-09-912-628-4
2	704	99.7	706	9	US-10-116-166-3
3	586	83.0	599	9	US-09-912-628-14
4	586	83.0	599	9	US-10-116-166-12
5	577	81.7	608	9	US-09-912-628-16
6	577	81.7	608	9	US-10-116-166-14
7	462.8	65.6	943	10	US-09-728-952-86
8	251.2	35.6	1624	10	US-09-925-300-500
9	246.8	35.0	1465	10	US-09-880-107-3027
10	212.8	30.1	1476	9	US-09-974-298-94
11	152.6	21.6	456	10	US-09-867-701-4671
12	150.6	21.3	484	9	US-09-918-995-2112
13	142.4	20.2	286	9	US-10-060-036-1797
14	99.8	14.1	290	9	US-10-025-380-288
15	99.8	14.1	290	10	US-09-815-343-1188
16	99.8	14.1	290	10	US-09-815-343-1348
17	99.8	14.1	290	10	US-09-922-217-288
18	99.8	14.1	290	10	US-09-833-263-288
19	88	12.5	1193	7	US-08-731-566-1

20 83.8 11.9 1931 10 US-09-925-301-358  
21 83.2 11.8 1284 10 US-09-954-456-1841  
22 77.4 11.0 1316 10 US-09-969-347-192  
23 75.6 10.7 156 10 US-09-998-598-2316  
24 73.6 10.4 953 7 US-08-731-566-3  
25 70 9.9 1950 9 US-10-091-442-30  
26 70 9.9 1950 10 US-09-140-719-30  
27 69.6 9.9 116 9 US-09-796-692-7379  
28 69.6 9.9 116 9 US-10-040-862-7379  
29 65.8 9.3 444 10 US-09-960-352-14649  
30 65.4 9.3 1308 10 US-09-735-705-109  
31 65.4 9.3 1308 10 US-09-850-716A-109  
32 65.4 9.3 1308 10 US-09-897-778-109  
33 65.4 9.3 1419 10 US-08-735-705-111  
34 65.4 9.3 1419 10 US-09-850-716A-111  
35 65.4 9.3 1419 10 US-09-897-778-111  
36 64.2 9.1 418 10 US-09-960-352-7066  
37 62.6 8.9 430 10 US-09-960-352-10531  
38 61.8 8.8 3279 9 US-10-012-896-382  
39 61.8 8.8 3279 9 US-09-895-793-382  
40 61.8 8.8 3279 9 US-09-895-814-382  
41 61.8 8.8 3279 9 US-10-010-940-382  
42 61.8 8.8 3279 10 US-09-759-143-382  
43 61.8 8.8 3279 10 US-09-780-669-382  
44 61.8 8.8 3279 10 US-09-822-827-382  
45 60.4 8.6 416 10 US-09-960-352-2546

ALIGNMENTS

RESULT 1

US-09-912-628-4

; Sequence 4, Application US/09912628

; Patent No. US20020160491A1

; GENERAL INFORMATION:

; APPLICANT: Ni et al.

; TITLE OF INVENTION: Human Serpin Polynucleotides, Polypeptides, and Antibodies

; FILE REFERENCE: PT00192

; CURRENT APPLICATION NUMBER: US/09/912,628

; CURRENT FILING DATE: 2001-07-26

; PRIOR APPLICATION NUMBER: PCT/US01/02484

; PRIOR FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 60/178,769

; PRIOR FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: PCT/US00/05082

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 706

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (635)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (655)

; OTHER INFORMATION: n equals a,t,g, or c

Query Match 99.7%; Score 704; DB 9; Length 706;

Best Local Similarity 100.0%; Pred. No. 8.5e-219;

Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACGAGCTTCGCTCTCGGCGAGCTGCGCGGAGAACTCGGCGTCCACCGTCATGATGCT 60

Db 1 GGACGAGCTTCGCTCTCGGCGAGCTGCGCGGAGAACTCGGCGTCCACCGTCATGATGCT 60

Qy 61 CTATCAGAACCAATGCGACATTTGCATTAAACCTTTTAAAAAGCTAGGGGAAAAACAC 120

Db 61 CTATCAGAACCAATGCGACATTTGCATTAAACCTTTTAAAAAGCTAGGGGAAAAACAC 120

LOCATION: (635)  
OTHER INFORMATION: n equals a,t,g, or c  
US-10-116-166-3

Query Match  
Best Local Similarity 99.7%; Score 704; DB 9; Length 706;  
Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGAGCTTCGCTCGCTGGCGAGCTCGCGGAGAACTGGGGCTCACCGTCATGATGCT 60  
DB 1 GGCAGAGCTTCGCTCGCTCGCTGGCGAGCTCGCGGAGAACTGGGGCTCACCGTCATGATGCT 60

QY 61 CTATCAGAAGCAAAATGCAATTTGCAATTAACCTTTTGAAGAGCTAGGGGAAAAACAAC 120  
DB 61 CTATCAGAAGCAAAATGCAATTTGCAATTAACCTTTTGAAGAGCTAGGGGAAAAACAAC 120

QY 121 TCAAAACAACCTTTATTTTCCCATGAGCATATCATCAGCTTGGCCATGTTTTCATGG 180  
DB 121 TCAAAACAACCTTTATTTTCCCATGAGCATATCATCAGCTTGGCCATGTTTTCATGG 180

QY 181 GGGCAAAAGGAAAAACATTCAGCTCAGATGCTCTCAGGCACTTTTGTAGTAAAAATCGAG 240  
DB 181 GGGCAAAAGGAAAAACATTCAGCTCAGATGCTCTCAGGCACTTTTGTAGTAAAAATCGAG 240

QY 241 GTGAAGATGGAGATATTCATCAGAGCTTTTCACTCTCTTGTGCAATTAACAGAACTG 300  
DB 241 GTGAAGATGGAGATATTCATCAGAGCTTTTCACTCTCTTGTGCAATTAACAGAACTG 300

QY 301 ACACCTGGAATATGCTTAGAAGCTCCCAACGGGCTCTTTGGAGAAAAGTCTTATGATTTCC 360  
DB 301 ACACCTGGAATATGCTTAGAAGCTCCCAACGGGCTCTTTGGAGAAAAGTCTTATGATTTCC 360

QY 361 TCACAGGTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCAACGATAAAACAGCTAGACT 420  
DB 361 TCACAGGTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCAACGATAAAACAGCTAGACT 420

QY 421 TTGTAATGATACAGAGAACTCCACACAGCTGTAACTCTCTGGTCTGATAAACTA 480  
DB 421 TTGTAATGATACAGAGAACTCCACACAGCTGTAACTCTCTGGTCTGATAAACTA 480

QY 481 AAGCCTGGAAAATTTATTCAAAACAGCTGTGCATCTCTGGAGAGCCAGGAATCGCCTTT 540  
DB 481 AAGCCTGGAAAATTTATTCAAAACAGCTGTGCATCTCTGGAGAGCCAGGAATCGCCTTT 540

QY 541 CCTCTGTGTACTGCAAGAGCTGCTTTCAGAGCCCTTTCAGAGCCCTTACTGTTTCCCAAT 600  
DB 541 CCTCTGTGTACTGCAAGAGCTGCTTTCAGAGCCCTTTCAGAGCCCTTACTGTTTCCCAAT 600

QY 601 GCAACTCTCTGTGACCCCGCATGCTGTGTGNCCTCCCTCCCTGTGAGCAGNTGTGA 660  
DB 601 GCAACTCTCTGTGACCCCGCATGCTGTGTGNCCTCCCTCCCTGTGAGCAGNTGTGA 660

QY 661 CTAATAAACTGCGGCAATTTTCATCTGTAAAAAATAAAAAAAAAAAAAA 706  
DB 661 CTAATAAACTGCGGCAATTTTCATCTGTAAAAAATAAAAAAAAAAAAAA 706

RESULT 3  
US-09-912-628-14  
Sequence 14, Application US/09912628  
Patent No. US20020160491A1  
GENERAL INFORMATION:  
APPLICANT: Ni et al.  
TITLE OF INVENTION: Human Serpin Polynucleotides, Polypeptides, and Antibodies  
FILE REFERENCE: PT001P2  
CURRENT APPLICATION NUMBER: US/09/912,628  
CURRENT FILING DATE: 2001-07-26  
PRIOR APPLICATION NUMBER: PCT/US01/02484  
PRIOR FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 60/178,769  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: PCT/US00/05082  
PRIOR FILING DATE: 2000-02-29

QY 121 TCAAAACAACCTTTATTTTCCCATGAGCATATCATCAGCTTGGCCATGTTTTCATGG 180  
DB 121 TCAAAACAACCTTTATTTTCCCATGAGCATATCATCAGCTTGGCCATGTTTTCATGG 180

QY 181 GGGCAAAAGGAAAAACATTCAGCTCAGATGCTCTCAGGCACTTTTGTAGTAAAAATCGAG 240  
DB 181 GGGCAAAAGGAAAAACATTCAGCTCAGATGCTCTCAGGCACTTTTGTAGTAAAAATCGAG 240

QY 241 GTGAAGATGGAGATATTCATCAGAGCTTTTCACTCTCTTGTGCAATTAACAGAACTG 300  
DB 241 GTGAAGATGGAGATATTCATCAGAGCTTTTCACTCTCTTGTGCAATTAACAGAACTG 300

QY 301 ACACCTGGAATATGCTTAGAAGCTCCCAACGGGCTCTTTGGAGAAAAGTCTTATGATTTCC 360  
DB 301 ACACCTGGAATATGCTTAGAAGCTCCCAACGGGCTCTTTGGAGAAAAGTCTTATGATTTCC 360

QY 361 TCACAGGTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCAACGATAAAACAGCTAGACT 420  
DB 361 TCACAGGTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCAACGATAAAACAGCTAGACT 420

QY 421 TTGTAATGATACAGAGAACTCCACACAGCTGTAACTCTCTGGTCTGATAAACTA 480  
DB 421 TTGTAATGATACAGAGAACTCCACACAGCTGTAACTCTCTGGTCTGATAAACTA 480

QY 481 AAGCCTGGAAAATTTATTCAAAACAGCTGTGCATCTCTGGAGAGCCAGGAATCGCCTTT 540  
DB 481 AAGCCTGGAAAATTTATTCAAAACAGCTGTGCATCTCTGGAGAGCCAGGAATCGCCTTT 540

QY 541 CCTCTGTGTACTGCAAGAGCTGCTTTCAGAGCCCTTTCAGAGCCCTTACTGTTTCCCAAT 600  
DB 541 CCTCTGTGTACTGCAAGAGCTGCTTTCAGAGCCCTTTCAGAGCCCTTACTGTTTCCCAAT 600

QY 601 GCAACTCTCTGTGACCCCGCATGCTGTGTGNCCTCCCTCCCTGTGAGCAGNTGTGA 660  
DB 601 GCAACTCTCTGTGACCCCGCATGCTGTGTGNCCTCCCTCCCTGTGAGCAGNTGTGA 660

QY 661 CTAATAAACTGCGGCAATTTTCATCTGTAAAAAATAAAAAAAAAAAAAA 706  
DB 661 CTAATAAACTGCGGCAATTTTCATCTGTAAAAAATAAAAAAAAAAAAAA 706

RESULT 2  
US-10-116-166-3  
Sequence 3, Application US/10116166  
Publication No. US20030040097A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: Serpin Polynucleotides, Polypeptides, and Antibodies  
FILE REFERENCE: PT001P1  
CURRENT APPLICATION NUMBER: US/10/116,166  
CURRENT FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: 09/641,721  
PRIOR FILING DATE: 2000-08-21  
PRIOR APPLICATION NUMBER: PCT/US00/05092  
PRIOR FILING DATE: 2000-02-29  
PRIOR APPLICATION NUMBER: 60/122,276  
PRIOR FILING DATE: 1999-03-01  
PRIOR APPLICATION NUMBER: 60/124,094  
PRIOR FILING DATE: 1999-03-12  
PRIOR APPLICATION NUMBER: 60/149,452  
PRIOR FILING DATE: 1999-08-18  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 706  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (635)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE

; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 599  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (585)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-912-628-14

Query Match 83.0%; Score 586; DB 9; Length 599;  
Best Local Similarity 99.8%; Pred. No. 2.2e-180;  
Matches 598; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 52 ATGGATGCTCTATCAGAGCAAAATGGCACATTTTGCATTAAACCTTTTGAAGAGCTAGGG 111  
1 ATGGATGCTCTATCAGAGCAAAATGGCACATTTTGCATTAAACCTTTTGAAGAGCTAGGG 60  
QY 112 GAAACAACTCAAACTTATTTT-CCCCATGAGCATATCATCAGCCTTTGGCCATG 170  
61 GAAACAACTCAAACTTATTTT-CCCCATGAGCATATCATCAGCCTTTGGCCATG 120  
QY 171 GTTTTCATGGGGCAAGGAAACACTGCAGCTCAGATGTCTCAGGCACTTTGTTTAGT 230  
121 GTTTTCATGGGGCAAGGAAACACTGCAGCTCAGATGTCTCAGGCACTTTGTTTAGT 180  
QY 231 AAAATCGAGGTGAAGATGAGATATTCATCGAGTTTTCAGTCACTTCTTGTGCAATT 290  
181 AAAATCGAGGTGAAGATGAGATATTCATCGAGTTTTCAGTCACTTCTTGTGCAATT 240  
QY 291 AACAGAACTGACACTGAATATGCTTAGAACTGCCAACGGGCTCTTTGGAGAAAAAGTCT 350  
241 AACAGAACTGACACTGAATATGCTTAGAACTGCCAACGGGCTCTTTGGAGAAAAAGTCT 300  
QY 351 TATGATTTCTCACAGTTTTACAGATTCCTGTGGCAAAATTCACCAAGCAACGATAAAA 410  
301 TATGATTTCTCACAGTTTTACAGATTCCTGTGGCAAAATTCACCAAGCAACGATAAAA 360  
QY 411 CAGTAGACTTTGTGAATGATACAGAGAGTCCACAAACAGCTGTAACTCCTGGGTTGCT 470  
361 CAGTAGACTTTGTGAATGATACAGAGAGTCCACAAACAGCTGTAACTCCTGGGTTGCT 420  
QY 471 GATAAACTAAAGCTTGGAAAAATTTTCAAAAGGCTGTCAATCTGGAGGAGCCAGGA 530  
421 GATAAACTAAAGCTTGGAAAAATTTTCAAAAGGCTGTCAATCTGGAGGAGCCAGGA 480  
QY 531 ATCGCCTCTTCTCTTGTACTGCAAGGCTGCTTTCAGAGCCCTACTGGTTCACTCT 590  
481 ATCGCCTCTTCTCTTGTACTGCAAGGCTGCTTTCAGAGCCCTACTGGTTCACTCT 540  
QY 591 ATTCCTCAATGCAACTCTCTGTGACCCCGCATGGCATGTGGTGNCCCTCCCTCTGTG 649  
541 ATTCCTCAATGCAACTCTCTGTGACCCCGCATGGCATGTGGTGNCCCTCCCTCTGTG 599

## RESULT 4

US-10-116-166-12  
; Sequence 12, Application US/10116166  
; Publication No. US20030040097A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Serpin Polynucleotides, Polypeptides, and Antibodies  
; FILE REFERENCE: PT001P1  
; CURRENT APPLICATION NUMBER: US/10/116,166  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 09/641,721  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: PCT/US00/05092  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: 60/122,276  
; PRIOR FILING DATE: 1999-03-01

; PRIOR APPLICATION NUMBER: 60/124,094  
; PRIOR FILING DATE: 1999-03-12  
; PRIOR APPLICATION NUMBER: 60/149,452  
; PRIOR FILING DATE: 1999-08-18  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 599  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (585)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-116-166-12

Query Match 83.0%; Score 586; DB 9; Length 599;  
Best Local Similarity 99.8%; Pred. No. 2.2e-180;  
Matches 598; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 52 ATGGATGCTCTATCAGAGCAAAATGGCACATTTTGCATTAAACCTTTTGAAGAGCTAGGG 111  
Db 1 ATGGATGCTCTATCAGAGCAAAATGGCACATTTTGCATTAAACCTTTTGAAGAGCTAGGG 60  
QY 112 GAAACAACTCAAACTTATTTT-CCCCATGAGCATATCATCAGCCTTTGGCCATG 170  
Db 61 GAAACAACTCAAACTTATTTT-CCCCATGAGCATATCATCAGCCTTTGGCCATG 120  
QY 171 GTTTTCATGGGGCAAGGAAACACTGCAGCTCAGATGTCTCAGGCACTTTGTTTAGT 230  
Db 121 GTTTTCATGGGGCAAGGAAACACTGCAGCTCAGATGTCTCAGGCACTTTGTTTAGT 180  
QY 231 AAAATCGAGGTGAAGATGAGATATTCATCGAGTTTTCAGTCACTTCTTGTGCAATT 290  
Db 181 AAAATCGAGGTGAAGATGAGATATTCATCGAGTTTTCAGTCACTTCTTGTGCAATT 240  
QY 291 AACAGAACTGACACTGAATATGCTTAGAACTGCCAACGGGCTCTTTGGAGAAAAAGTCT 350  
Db 241 AACAGAACTGACACTGAATATGCTTAGAACTGCCAACGGGCTCTTTGGAGAAAAAGTCT 300  
QY 351 TATGATTTCTCACAGTTTTACAGATTCCTGTGGCAAAATTCACCAAGCAACGATAAAA 410  
Db 301 TATGATTTCTCACAGTTTTACAGATTCCTGTGGCAAAATTCACCAAGCAACGATAAAA 360  
QY 411 CAGTAGACTTTGTGAATGATACAGAGAGTCCACAAACAGCTGTAACTCCTGGGTTGCT 470  
Db 361 CAGTAGACTTTGTGAATGATACAGAGAGTCCACAAACAGCTGTAACTCCTGGGTTGCT 420  
QY 471 GATAAACTAAAGCTTGGAAAAATTTTCAAAAGGCTGTCAATCTGGAGGAGCCAGGA 530  
Db 421 GATAAACTAAAGCTTGGAAAAATTTTCAAAAGGCTGTCAATCTGGAGGAGCCAGGA 480  
QY 531 ATCGCCTCTTCTCTTGTACTGCAAGGCTGCTTTCAGAGCCCTACTGGTTCACTCT 590  
Db 481 ATCGCCTCTTCTCTTGTACTGCAAGGCTGCTTTCAGAGCCCTACTGGTTCACTCT 540  
QY 591 ATTCCTCAATGCAACTCTCTGTGACCCCGCATGGCATGTGGTGNCCCTCCCTCTGTG 649  
Db 541 ATTCCTCAATGCAACTCTCTGTGACCCCGCATGGCATGTGGTGNCCCTCCCTCTGTG 599

## RESULT 5

US-09-912-628-16  
; Sequence 16, Application US/09912628  
; Patent No. US20020160491A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: Human Serpin Polynucleotides, Polypeptides, and Antibodies  
; FILE REFERENCE: PT001P2  
; CURRENT APPLICATION NUMBER: US/09/912,628  
; CURRENT FILING DATE: 2001-07-26  
; PRIOR APPLICATION NUMBER: PCT/US01/02484  
; PRIOR FILING DATE: 2001-01-26

	PRIOR APPLICATION NUMBER: 60/178,769	
	PRIOR FILING DATE: 2000-01-28	
	PRIOR APPLICATION NUMBER: PCT/US00/05082	
	PRIOR FILING DATE: 2000-02-29	
	NUMBER OF SEQ ID NOS: 17	
	SOFTWARE: PatentIn Ver. 2.0	
	SEQ ID NO 16	
	LENGTH: 608	
	TYPE: DNA	
	ORGANISM: Homo sapiens	
	FEATURE:	
	NAME/KEY: SITE	
	LOCATION: (91)	
	OTHER INFORMATION: n equals a,t,g, or c	
	NAME/KEY: SITE	
	LOCATION: (92)	
	OTHER INFORMATION: n equals a,t,g, or c	
	NAME/KEY: SITE	
	LOCATION: (93)	
	OTHER INFORMATION: n equals a,t,g, or c	
	NAME/KEY: SITE	
	LOCATION: (94)	
	OTHER INFORMATION: n equals a,t,g, or c	
	NAME/KEY: SITE	
	LOCATION: (95)	
	OTHER INFORMATION: n equals a,t,g, or c	
	NAME/KEY: SITE	
	LOCATION: (96)	
	OTHER INFORMATION: n equals a,t,g, or c	
	NAME/KEY: SITE	
	LOCATION: (97)	
	OTHER INFORMATION: n equals a,t,g, or c	
	NAME/KEY: SITE	
	LOCATION: (98)	
	OTHER INFORMATION: n equals a,t,g, or c	
	NAME/KEY: SITE	
	LOCATION: (99)	
	OTHER INFORMATION: n equals a,t,g, or c	
	NAME/KEY: SITE	
	LOCATION: (594)	
	OTHER INFORMATION: n equals a,t,g, or c	
	US-09-912-628-16	
	Query Match	81.7%; Score 577; DB 9; Length 608;
	Best Local Similarity	98.4%; Pred.No. 1.9e-177;
	Matches 598; Conservative	0; Mismatches 0; Indels 10; Gaps 1;
QY	52 ATGGATGCTCTATCAGAAGCAAAATGGACATTTCGATTAAACCTTTTGA AAAAGCTAGGG 111	
D	1 ATGGATGCTCTATCAGAAGCAAAATGGACATTTCGATTAAACCTTTTGA AAAAGCTAGGG 60	
DB	112 GAAAACAACCTCAAACAACCTATTTTTTTC-----CCCATGAGCATATCATGACC 161	
	61 GAAAAACAACCTCAAACAACCTATTTTTTTCNNNNNNNCCCATGAGCATATCATGACC 120	
QY	162 TTGGCCATGTTTTCATGGGGCAAGGAAACACTCGAGCTCAGATGTCTCAGGCACCTT 221	
DB	121 TTGGCCATGTTTTCATGGGGCAAGGAAACACTCGAGCTCAGATGTCTCAGGCACCTT 180	
QY	222 TGTTTTAGTAAATCGGAGGTGAAGTGAGATATTCATCAGGTTTTCAGTCACTCTTT 281	
DB	181 TGTTTTAGTAAATCGGAGGTGAAGTGAGATATTCATCAGGTTTTCAGTCACTCTTT 240	
QY	282 GTTGCAATTAACAGAACTCACACTGAATATGTCTTAGAATGCCAACGGGCTCTTTGGA 341	
DB	241 GTTGCAATTAACAGAACTCACACTGAATATGTCTTAGAATGCCAACGGGCTCTTTGGA 300	
QY	342 GAAAGCTCTTATGATTTCTCTACAGTTTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCA 401	
DB	301 GAAAGCTCTTATGATTTCTCTACAGTTTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCA 360	
QY	402 ACGATAAAACAGCTAGCTTTTGTGAATGATACAGAGAAAGTCCACAACACGCTGTAACCTCC 461	

Db	361 ACGATAAAACAGCTAGCTTTTGTGAATGATACAGAGAAAGTCCACAACACGCTGTAACCTCC 420
QY	462 TGGTTTGTCTGATAAACCTAAAGCCTGGAAAAATTTTCAAAACAGCCTGTGCATCTGGAG 521
Db	421 TGGTTTGTCTGATAAACCTAAAGCCTGGAAAAATTTTCAAAACAGCCTGTGCATCTGGAG 480
QY	522 GAGCCAGGAATCGCCTCTTCTCTTGTGTACTGCAAGCCCTGCTTTTACAGAGCCCCCTACTG 581
Db	481 GAGCCAGGAATCGCCTCTTCTCTTGTGTACTGCAAGCCCTGCTTTTACAGAGCCCCCTACTG 540
QY	582 GTTCACCTCTATTTCCTCAAAATGCAACTCTCTCTGTGACCCCCTGATGGCATGTGTGNCTCC 641
Db	541 GTTCACCTCTATTTCCTCAAAATGCAACTCTCTCTGTGACCCCCTGATGGCATGTGTGNCTCC 600
QY	642 TCCCTGTG 649
Db	601 TCCCTGTG 608
	RESULT 6
	US-10-116-166-14
	; Sequence 14, Application US/10116166
	; Publication No. US20030040097A1
	; GENERAL INFORMATION:
	; APPLICANT: Ruben et al.
	; TITLE OF INVENTION: Serpin Polynucleotides, Polypeptides, and Antibodies
	; FILE REFERENCE: PT001P1
	; CURRENT APPLICATION NUMBER: US/10/116,166
	; CURRENT FILING DATE: 2002-04-05
	; PRIOR APPLICATION NUMBER: 09/641,721
	; PRIOR FILING DATE: 2000-08-21
	; PRIOR APPLICATION NUMBER: PCT/US00/05092
	; PRIOR FILING DATE: 2000-02-29
	; PRIOR APPLICATION NUMBER: 60/122,276
	; PRIOR FILING DATE: 1999-03-01
	; PRIOR APPLICATION NUMBER: 60/124,094
	; PRIOR FILING DATE: 1999-03-12
	; PRIOR APPLICATION NUMBER: 60/149,452
	; PRIOR FILING DATE: 1999-08-18
	; NUMBER OF SEQ ID NOS: 15
	; SOFTWARE: PatentIn Ver. 2.0
	; SEQ ID NO 14
	; LENGTH: 608
	; TYPE: DNA
	; ORGANISM: Homo sapiens
	; FEATURE:
	; NAME/KEY: SITE
	LOCATION: (91)
	OTHER INFORMATION: n equals a,t,g, or c
	NAME/KEY: SITE
	LOCATION: (92)
	OTHER INFORMATION: n equals a,t,g, or c
	NAME/KEY: SITE
	LOCATION: (93)
	OTHER INFORMATION: n equals a,t,g, or c
	NAME/KEY: SITE
	LOCATION: (94)
	OTHER INFORMATION: n equals a,t,g, or c
	NAME/KEY: SITE
	LOCATION: (95)
	OTHER INFORMATION: n equals a,t,g, or c
	NAME/KEY: SITE
	LOCATION: (96)
	OTHER INFORMATION: n equals a,t,g, or c
	NAME/KEY: SITE
	LOCATION: (97)
	OTHER INFORMATION: n equals a,t,g, or c
	NAME/KEY: SITE
	LOCATION: (98)
	OTHER INFORMATION: n equals a,t,g, or c
	NAME/KEY: SITE
	LOCATION: (99)
	OTHER INFORMATION: n equals a,t,g, or c

```

; NAME/KEY: SITE
; LOCATION: (594)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-116-166-14

Query Match
Best Local Similarity 81.7%; Score 577; DB 9; Length 608;
Matches 598; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

Qy 52 ATGGATGCTCTATCAGAGCAAAATGGACATTTTCATTAACCTTTTGAAAAAGCTAGGG 111
Db 1 ATGGATGCTCTATCAGAGCAAAATGGACATTTTCATTAACCTTTTGAAAAAGCTAGGG 60

Qy 112 GAAACAACTCAAACTTATTTTTC-----CCCATGAGCATATCATCAGCC 161
Db 61 GAAACAACTCAAACTTATTTTTCNNNNNNNNCCATGAGCATATCATCAGCC 120

Qy 162 TTGGCCATGTTTTCATGGGGGCAAGGGAACACTGCAGCTCAGATGCTCAGGCACATT 221
Db 121 TTGGCCATGTTTTCATGGGGGCAAGGGAACACTGCAGCTCAGATGCTCAGGCACATT 180

Qy 222 TGTTTTAGTAAATCGAGGTGAAGATGGAGATATTTCAGAGTTTTCAGTCACCTTCTT 281
Db 181 TGTTTTAGTAAATCGAGGTGAAGATGGAGATATTTCAGAGTTTTCAGTCACCTTCTT 240

Qy 282 GTTGCAATTAAACAGAACTGACACTGAATATGCTTAGAACTGCCAAGGGCTCTTTGGA 341
Db 241 GTTGCAATTAAACAGAACTGACACTGAATATGCTTAGAACTGCCAAGGGCTCTTTGGA 300

Qy 342 GAAAAGTCTTATGATTTCTCTCAGAGTTTACAGATTTCTGTAAGTCCAACTCC 461
Db 301 GAAAAGTCTTATGATTTCTCTCAGAGTTTACAGATTTCTGTAAGTCCAACTCC 420

Qy 462 TGGGTGTGTAATAAACTAAAGCTTGGAAATTTATCAACAGAGCTGTACATCTGGAG 521
Db 421 TGGGTGTGTAATAAACTAAAGCTTGGAAATTTATCAACAGAGCTGTACATCTGGAG 480

Qy 522 GAGCCAGGAATCGCTCTCTCTTGTGTAAGCTGCTGCTTTTCAAGCCCTACTG 581
Db 481 GAGCCAGGAATCGCTCTCTCTTGTGTAAGCTGCTGCTTTTCAAGCCCTACTG 540

Qy 582 GTTCACTCTATTTCCAAATCAACTCTCTCTGTGACCCCGCATGCGATGGTGNCTCCC 641
Db 541 GTTCACTCTATTTCCAAATCAACTCTCTCTGTGACCCCGCATGCGATGGTGNCTCCC 600

Qy 642 TCCCTGTG 649
Db 601 TCCCTGTG 608

```

## RESULT 7

```

US-09-728-952-86
; Sequence 86, Application US/09728952
; Patent No. US20020111302A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Vicki
; APPLICANT: Ujwal, Manusha L.
; TITLE OF INVENTION: No. US20020111302alel Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 799
; CURRENT APPLICATION NUMBER: US/09/728,952

```

```

; CURRENT FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: pt_FL_Genes Version 2.0
; SEQ ID NO 86
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (189)..(701)
US-09-728-952-86

```

```

Query Match
Best Local Similarity 65.6%; Score 462.8; DB 10; Length 943;
Matches 478; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 9 CTTGCGCTCTGGGCGAGTGGGGAGAACTGGGGCTCACCGTCATGGATGCTCTATCAGA 68
Db 146 CTTGCGCTCTGGGCGAGTGGGGAGAACTGGGGCTCACCGTCATGGATGCTCTATCAGA 205

Qy 69 AGCAAAATGGCACATTTTGCATTAAACCTTTTGAAGAAAGCTAGGGGAAAAACAACCAAA 128
Db 206 AGCAAAATGGCACATTTTGCATTAAACCTTTTGAAGAAAGCTAGGGGAAAAACAACCAAA 265

Qy 129 CTTATTTTTCCTCCCA-IGAGCATATCATCAGCATTTGTTTATGTAAGTAAATCGGAGTGAAGA 247
Db 266 CTTATTTTTCCTCCCACTGAGCATATCATCAGCATTTGTTTATGTAAGTAAATCGGAGTGAAGA 385

Qy 188 GGGAAACACTGCAGCTCAGATGCTCAGGCACTTTGTTTATGTAAGTAAATCGGAGTGAAGA 247
Db 326 GGGAAACACTGCAGCTCAGATGCTCAGGCACTTTGTTTATGTAAGTAAATCGGAGTGAAGA 307

Qy 248 TGGAGATATTTCATCGAGGTTTTCAGTCACCTTTGTTGCAATTAACAGAACTGACACTGA 445
Db 386 TGGAGATATTTCATCGAGGTTTTCAGTCACCTTTGTTGCAATTAACAGAACTGACACTGA 445

Qy 308 ATATGTGCTTAGAACTGCCAAGCGGCTCTTTGGAGAAAAAGCTTTATGATTTCCTCACAGG 367
Db 446 ATATGTGCTTAGAACTGCCAAGCGGCTCTTTGGAGAAAAAGCTTTATGATTTCCTCACAGG 305

Qy 368 TTTTACAGATTCCTGTGGCAAAATTTCTACCAAGCAACGATAAACAGCTAGACTTTGTGAA 427
Db 506 TTTTACAGATTCCTGTGGCAAAATTTTACCAAGCAACGATAAACAGCTAGACTTTGTGAA 565

Qy 428 TGATACAGAGAACTCCACAACTGTAAGTAACTCCTGGGTTGCTGTGATAAACTAAAGCTG 487
Db 566 TGATACAGAGAACTCCACAACTGTAAGTAACTCCTGGGTTGCTGTGATAAACTAAAGCTG 487

Qy 488 GAAAAAT 493
Db 626 AAAAAAT 631

```

## RESULT 8

```

US-09-925-300-500
; Sequence 500, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 500
; LENGTH: 1624
; TYPE: DNA

```

RESULT 9  
US-09-880-107-3027  
; Sequence 3027, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3027  
; LENGTH: 1465  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 S63272  
US-09-880-107-3027  
Query Match 35.0%; Score 246.8; DB 10; Length 1465;  
Best Local Similarity 75.8%; Pred. No. 1.5e-69;  
Matches 33%; Conservative 0; Mismatches 102; Indels 4; Gaps 106  
47 CCTCATGGATGCTCTATCAGCAAGCAAAATGGCACATTTCGACATTAAACCTTTGAAAAAGC 106

```

; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 94
; LENGTH: 1476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 589880CB1
; US-09-974-298-94

Query Match      30.1%; Score 212.8; DB 9; Length 1476;
Best Local Similarity 73.1%; Pred. No. 1.8e-58;
Matches 320; Conservative 0; Mismatches 102; Indels 16; Gaps 3;

      QY      50  TCATGGATGCTCTATCAGACCAATGGCACATTTGCATTAAACCTTTTCGAAAAAGCTAG 109
      DB      108  TGAATGGATGACCTCTGTGAACCAATGGCACATTTTTCATCAGCTTATTAAATAATTGG 167

      QY      110  GGGAAA---ACAACTCAAACAACTATTATTTTTTC-CCATGAGCATATCATCAGCCTTGG 165
      DB      168  GGGAGGAGGACAACTCAAGAACGTATTCTCTCCCATGAGCATCTCCTGCCCCCTGG 227

      QY      166  CCATGGTTTTTCATGGGGCAAGGGAACACTGGAGGTCAGATGCTCTCAGGCACTTTGTT 225
      DB      228  CCATGGTCTTTCATGGGGCAAGGGAAGACATGACGCCACATGTCCTCCAGGCACTTTGTT 287

      QY      226  TTAGTAAAAATCGGAGGTGAAGATGGAGATATTTCATCGAGGTTTTTCAGTCACCTCTTGTG 285
      DB      288  T-----ATACAAGACGGAGATATTCCACCGAGGTTTCCAGTCACCTTCTTCAGTG 335

```

QY	286	CAATTAAACGAAC	TGACACTGAC	TAAATATG	TGCTTAGAAC	TGCCAACGGG	CTCTTTGGAGAA	345
Db	336	AAGTTAAACGAAC	TGGCACTCAG	TACTTGGCT	TTAGAACTGC	CAACAGACT	CTTTGGAGAA	395
QY	346	AGTCTTATGAT	TTCCCTCAGAG	TTTTACAG	ATTCTCTGGG	AAATTTCTTAC	CAAGCAACGA	405
Db	396	AGACGTGTG	ATTTCCCTTC	CAGACTTT	TAAGAATACT	GTGCAAGT	TTCTATCAGG	455
QY	406	TAAACAGCT	TAGACTTTG	TGAAATG	TACAGAGAAG	TCCACAAACAG	TGTAAACTCTCTGGG	465
Db	456	TGGAGGAG	TGTGCTT	TGCTGAAG	CACACTGAAG	AGAGTGCAGG	AGCATATAAAT	515
QY	466	TTGCTGAT	AAAACTAAAG	483				
Db	516	TGCGACAG	AAGACTGAAG	533				

.T 11  
 US-09-867-701-4671  
 ; Sequence 4671, Application US/09867701  
 ; Patent No. US2002013237A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Aglate, Paul A.  
 ; APPLICANT: Jones, Robert  
 ; APPLICANT: Harlocker, Susan L.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
 ; FILE REFERENCE: 210121.497  
 ; CURRENT APPLICATION NUMBER: US/09/867,701  
 ; CURRENT FILING DATE: 2001-05-29  
 ; NUMBER OF SEQ ID NOS: 10912  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4671  
 ; LENGTH: 456  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 ; US-09-867-701-4671

Query Match	21.6%;	Score 152.6;	DB 10;	Length 456;
Best Local Similarity	67.0%;	Pred. No. 3.4e-39;		
Matches 254;	Conservative 0;	Mismatches 109;	Indels 16;	Gaps 12;
QY	41	GGCTCACCGTCATCGATGCTCTATCAGAACCAATGGCAATTTGGCATTAACCTTTTGA	100	
Db	79	GGCCCTGCATCATGGAAACTCTTTCTTAATGCAAGTGTGTACTTTTGGCCATAGCGCTTTAA	138	
QY	101	AAAAGCTAGGGGAAAAACAACCTCAAAACAAC-----TTATTTTTTTTCCCATGAGCATATCAT	156	
Db	139	AGATACTGTGTCAAGATAACCCCTTCGCACAACTGTGTCTCTCTGTGAGCATCTCCT	198	
QY	157	CAGCCTTGGCCATGGTTTTTCATCGGGGCAAGGAAACAATCGAGCTCAGATGTCTCAGG	216	
Db	199	CTGCCCTGGCCATGGTTTTCTCTTAGGGGCAAGGAAACAACCGCAACCCAGATGCCCCAGG	258	
QY	217	CACCTTGTGTTTAGTAAAAATCGGAGGTGAAGATGGAGATATTATCATCGAGGTTTTCAAGTCAC	276	
Db	259	CACGTCTTTAAACA-----CAGAGGAAGACATTCATCGGGCTTTTCAGTCCG	306	
QY	277	TTCTTGTGTCAATTAACAGAACTGAACATGGAATATGTGCTTAGAACTGSCCAACGGGCTCT	336	
Db	307	TTCTACTGAAGTGAACAAGGCTGGGCACACAGTACTGCTGAGAACGGSCCAACAGGCTCT	366	
QY	337	TTGGAGAAAAAGTCTTATGATTTTCTCACAGGTTTTTACAGATTTCTGTGGGCAAAATCTTACC	396	
Db	367	TTGGAGAGAAAACCTTGTCAGTTCCTCTCAACGTTTTAAGGAATCCTGTCTTCAATTTCTACC	426	
QY	397	AAGCAACGATAAAACAGCT	415	
Db	427	ATGCTGAACCTGAAAGAGCT	445	

RESULT 12

```

US-09-918-995-2112
; Sequence 2112, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2112
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(484)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-2112

Query Match      21.3%; Score 150.6; DB 9; Length 484;
Best Local Similarity 67.0%; Pred. No. 1.6e-38;
Matches 235; Conservative 0; Mismatches 104; Indels 12; Gaps 1

Qy      133 TTTTTCCTCCCATCAGCATATCATCAGCCCTGGCCATGGTTTTCATCGGGGCAAAAGGAA 192
Db      56 TCTGTTCTCTCTGTGAGCATCTCTCTGCCCTGGCCATGGTTCTCTAGGGGCAAAAGGAA 115

Qy      193 ACATCTGAGCTCAGATGTCTCTCAGGCACCTTGTGTTTGTAGTAAATCGGAGGTGAGATCGAG 252
Db      116 ACACCGCAACCGAGATGGCCAGGCACCTGCTTTAAACA-----CAGAGGAAG 163

Qy      253 ATATTATCAGAGTTTTCAGTCACTTCTTGTGTGCAATTAACAGAACTGCACACTGAATATG 312
Db      164 ACATTTCATCGGGCTTTCCAGTCGCTTCTCACTGAACTGAACAAGGCTGGGCACACAGTACC 223

Qy      313 TGGCTTAGAAGTCGCCACGGGCTCTTTGGAGAAAAGTCTTATGATTTCTCTCAGAGTTTTTA 372
Db      224 TGGCTGAGAACGGTCAACAGGCTCTTTGGAGAGAAAACCTTGTTCAGTTCTCTCAACGTTTA 283

Qy      373 CAGATTCTCTGTGGCAAAATCTTACCAAGCAACGATAAAACAGCTTAGACTTTTGTGAATGATA 432
Db      284 AGGAATCTCTGTCTTCAATTCTACCATCTGAGCTGAGAGAGCTTTCCTTTATCAGAGCTG 343

Qy      433 CAGAGAAGTCCACAAACAGCTGTAAACTCTCTGGGTTGCTGTATAAACTAAAG 483
Db      344 CAGAAGAGTGCAGGAAACACATCAACACTCTGGTCTCAAAAAGACCTCGAG 394

```

```

RESULT 13
US-10-060-036-1797
; Sequence 1797, Application US10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060.036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FASTSEQ FOR WINDOWS Version 4.0
; SEQ ID NO 1797
; LENGTH: 286

```

Db 230 AAAAAATCTTACCAAGCAGAGATGGAGGAGCTTGACTTTTATCAGCGCGGTAGAGAAGTCCA 171  
Qy 445 CAACACGTGTAAACTCTCTGGTTCCTCATATAAACTAAAG 483  
Db 170 GAAACACATATAAACACCTGGGTAGCTGAAAGACAGAG 132

RESULT 15

US-09-815-343-1188  
; Sequence 1188, Application US/09815343  
; Patent No. US2001005596A1  
; GENERAL INFORMATION:  
; APPLICANT: Meagher, Madeleine  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: King, Gordon E.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; FILE REFERENCE: 210121.504  
; CURRENT APPLICATION NUMBER: US/09/815,343  
; CURRENT FILING DATE: 2001-03-22  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1188  
; LENGTH: 290  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-815-343-1188

Query Match 14.1%; Score 99.8; DB 10; Length 290;  
Best Local Similarity 76.7%; Pred. No. 4e-22;  
Matches 122; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 325 CCAACGGGCTCTTTGGAGAAAAGTCTTATGATTTCTCCTCAGAGTTTACAGATTCTCTGTG 384  
Db 1 CCAACAGGCTCTTTGGGAAAAGTCTTGTGATTTCTCTCATCTTTTAGAGATTCTCTGCC 60  
Qy 385 GCAATTTCTACCAACGACGATATAACAGCTAGACTTTGTGAATGATACAGAGAAGTCCA 444  
Db 61 AAAAAATTTACCAACGACGATGGAGGAGCTTTGACTTTATCAGCGCCGTAGAGAAGTCCA 120  
Qy 445 CAACACGTGTAAACTCTCTGGGTTGCTGATAAACTTAAAG 483  
Db 121 GAAACACATATAAACCTGGGTAGCTGAAAGACAGAGAAG 159

Search completed: May 22, 2003, 00:20:00  
Job time : 116 secs

; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-060-036-1797  
Query Match 20.2%; Score 142.4; DB 9; Length 286;  
Best Local Similarity 73.6%; Pred. No. 5.2e-36;  
Matches 209; Conservative 0; Mismatches 71; Indels 4; Gaps 2;  
Qy 27 GCGCGGAGAACTGGGGCTCACCGCTCATGGATGCTTATCAGAAAGCAAAATGGCACATTTGC 86  
Db 6 GCTCCCGCTCTGGAGTCTGCCATCATGGATGTTCTCGGAGAAAGCAAAATGGCACATTTGC 65  
Qy 87 ATTAACCTTTGAAAAGCTAGGGGAAAACAACTCAAACTATTTTTC-CCCAT 145  
Db 66 CTTAAACCTTTGAAAACGCTGGGTAAGACAACTCGAAGATGTGTTTCTCACCCAT 125  
Qy 146 GAGCATATCATAGCTTGGCCATGTTTTCATGGGGGCAAGGAAACACTGCAGCTCA 205  
Db 126 GAGCATGCTCTGGCCCTGGCCATGGTCTACATGGGGGCAAGGAAACACCGCTGCACA 185  
Qy 206 GATGCTCTCAGGCACTTTGTTTATAGTAAATTCGAGGTGAAGATGGAGATTTTCATCGAGG 265  
Db 186 GATGGCCAGACTTTCTTTCAATAAAAGTGGCGGT---GGTGGAGACATCCACCAGG 242  
Qy 266 TTTTCAGTCACTCTTGTGTCATTAACAGAACTGCAGCTGAAT 309  
Db 243 CTTCCAGTCTCTTCTCACCAGTGAACAGGCTGGCACCAGT 286

RESULT 14

US-10-025-380-288/c  
; Sequence 288, Application US/10025380  
; Publication No. US20020182191A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Stolk, John A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Smith, Carole L.  
; APPLICANT: King, Gordon E.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Clapper, Jonathan D.  
; APPLICANT: Skeiky, Yasir A. W.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick Thomas S.  
; APPLICANT: Carter, Darrick  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
; FILE REFERENCE: 210121.471C14  
; CURRENT APPLICATION NUMBER: US/10/025,380  
; CURRENT FILING DATE: 2001-12-19  
; NUMBER OF SEQ ID NOS: 1129  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 288  
; LENGTH: 290  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-025-380-288

Query Match 14.1%; Score 99.8; DB 9; Length 290;  
Best Local Similarity 76.7%; Pred. No. 4e-22;  
Matches 122; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 325 CCAACGGGCTCTTTGGAGAAAAGTCTTATGATTTCTCCTCAGAGTTTACAGATTCTCTGTG 384  
Db 290 CCAACAGGCTCTTTGGGAAAAGTCTTGTGATTTCTCTCATCTTTTAGAGATTCTCTGCC 231  
Qy 385 GCAATTTCTACCAACGACGATATAACAGCTAGACTTTGTGAATGATACAGAGAAGTCCA 444



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 21, 2003, 13:40:58 ; Search time 1505 Seconds  
(without alignments)  
7597.351 Million cell updates/sec

Title: US-09-912-628-4  
Perfect score: 706  
Sequence: 1 gscagagcttcgctcctgg.....gtaaaaaaaaaaaaaaa 706

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

hed: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estcov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	gb_gss:*
18:	em_gss_hum:*
19:	em_gss_inv:*
20:	em_gss_pln:*
21:	em_gss_vrt:*
22:	em_gss_fun:*
23:	em_gss_mam:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	475.4	67.3	618	14	BM696260
2	404.8	57.3	557	10	AW205887
3	274	38.8	1150	14	BM910193
4	253.2	35.9	788	13	BI603219
5	252.6	35.8	664	13	BI669663
6	252.6	35.8	804	12	BG749304

7	252.6	35.8	905	14	BQ681973
8	252	35.7	595	12	BG894979
9	251.6	35.6	536	10	BE543971
10	251.6	35.6	818	13	BI222840
11	251.6	35.6	875	12	BG108578
12	251.6	35.6	880	13	BM461676
13	251.6	35.6	927	14	BQ949817
14	251.6	35.6	957	13	BI822888
15	251.6	35.6	991	14	BQ072385
16	250.8	35.5	767	9	AI557224
17	250.4	35.5	573	13	BI772650
18	250.4	35.5	615	13	BM272500
19	250.4	35.5	743	13	BI222593
20	250.4	35.5	853	13	BI259189
21	250	35.4	662	10	BE384168
22	250	35.4	705	12	BG746783
23	250	35.4	731	12	BG477130
24	250	35.4	809	13	BI669865
25	250	35.4	835	12	BG540177
26	250	35.4	865	12	BG334413
27	250	35.4	928	14	BQ708836
28	250	35.4	939	14	BQ674001
29	250	35.4	969	14	BQ673166
30	250	35.4	1002	13	BM545433
31	250	35.4	1066	13	BI601472
32	249.6	35.4	649	12	BG437282
33	249.6	35.4	1098	14	BM913358
34	249.4	35.3	802	13	BI195978
35	249	35.3	731	9	AI119612
36	248.8	35.2	527	9	AI117472
37	248.8	35.2	651	13	BM503457
38	248.8	35.2	655	13	BM043698
39	248.8	35.2	897	14	BQ440763
40	248.4	35.2	795	12	BG216036
41	248.4	35.2	805	12	BG212435
42	248.4	35.2	814	12	BG197328
43	248.4	35.2	815	12	BG203037
44	248.4	35.2	872	12	BI81795
45	248.4	35.2	892	12	BG186517

ALIGNMENTS

RESULT 1  
BM696260  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

BM696260  
UI-E-DW0-agl-a-11-0-UI.r1  
BM696260  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 618)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

618 bp mRNA linear EST 28-FEB-2002  
UI-E-DW0-agl-a-11-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone  
mRNA sequence.

Clone Distribution: Researchers may obtain clones from Research Genetics (www.regen.com).

Seq primer: M13 Reverse.

# FEATURES

source  
1. .618  
Location/Qualifiers  
/organism="Homo sapiens"  
/db xref="taxon:9606"  
/clone="UI-E-DW0-agl-a-11-0-UI"  
/clone\_lib="UI-E-DW0"  
/tissue\_type="lens"  
/dev\_stage="adult"  
/lab host="PH10B (Life Technologies) (T1 phage resistant)"  
/notes="Organ: eye; Vector: p773-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-DW0 is a cDNA library containing the following tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (gt)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."  
BASE COUNT 192 a 125 c 121 g 180 t  
ORIGIN

Query Match 67.3%; Score 475.4; DB 14; Length 618;  
Best Local Similarity 98.8%; Pred. No. 5:5e-94; Indels 0; Gaps 0;  
Matches 479; Conservative 0; Mismatches 6;  
9 CTTGCTCTCTGGGAGCTGCGGAGAACTGGGCTCACCGTCATGATGCTCTATCAGA 68  
15 CTTGCTCTCTGGGAGCTGCGGAGAACTGGGCTCACCGTCATGATGCTCTATCAGA 74  
69 AGCAATGGCACTTTGATTAACCTTTTGAAGAGCTAGGGGAAACAACTCAACAA 128  
75 AGCAATGGCACTTTGATTAACCTTTTGAAGAGCTAGGGGAAACAACTCAACAA 134  
129 CTTATTTTTCCTCCATGAGCATATCATCAGCTTGGCCATGGTTTTCATGGGGCAAAG 188  
135 CTTATTTTTCCTCCATGAGCATATCATCAGCTTGGCCATGGTTTTCATGGGGCAAAG 194  
189 GGAAACACTGACGCTCAGATGCTCAGGACCTTTGTTAGTAAATCGAGGTGAAGAT 248  
195 GGAAACACTGACGCTCAGATGCTCAGGACCTTTGTTAGTAAATCGAGGTGAAGAT 254  
249 GGAGATATTCATCGAGGTTTTCAGTCACCTTCTTGTGCAATTAACAGAACTGACACTGAA 308  
255 GGAGATATTCATCGAGGTTTTCAGTCACCTTCTTGTGCAATTAACAGAACTGACACTGAA 314  
309 TATGTCCTTAGAAGTCCCAACGGGCTTTTGGAGAAAAGTCTTATGATTTCTCACAGGT 368  
315 TATGTCCTTAGAAGTCCCAACGGGCTTTTGGAGAAAAGTCTTATGATTTCTCACAGGT 374  
369 TTTACAGATTCCTGGGCAATTTCTACCAAGCAACGATAAATACAGCTAGACTTTGTGAAT 428  
375 TTTACAGATTCCTGGGCAATTTCTACCAAGCAACGATAAATACAGCTAGACTTTGTGAAT 434  
429 GATACAGAGAGTCCCAACACGCTGTAATCTCTGGTGTGCTGATGATAAATACAGCTGG 488  
435 GATACAGAGAGTCCCAACACGCTGTAATCTCTGGTGTGCTGATGATAAATACAGCTGG 494  
489 AAAAT 493  
495 AATAT 499

RESULT 2  
AW205887/c

## LOCUS

DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AW205887 557 bp mRNA linear EST 02-DEC-1999  
UI-H-B11-afw-d-02-0-UI-s1 NCI CGAP\_Sub3 Homo sapiens cDNA clone  
IMAGE:2723354 3', mRNA sequence.  
AW205887  
AW205887.1 GI:6505361  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 557)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
NCI-CGAP clone distribution information can be found through the  
I.M.A.G.E. Consortium/LLNL at:  
www.bio.lnl.gov/bbrp/image/image.html  
Seq primer: M13 Forward  
POLYA=Yes.

## FEATURES

Location/Qualifiers  
1. .557  
/organism="Homo sapiens"  
/db xref="taxon:9606"  
/clone="IMAGE:2723354"  
/clone\_lib="NCI CGAP\_Sub3"  
/lab host="PH10B (Life Technologies)"  
/note="vector: p773-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NCI CGAP Sub3 library is a subtracted library derived from the NCI CGAP Sub1 library, which is a subtracted library derived from BI. BI constitutes a mixture of 21 normalized or subtracted NCI CGAP libraries: NCI CGAP\_Co4, NCI CGAP\_Pr22, NCI CGAP\_Pr28, NCI CGAP\_Co10, NCI CGAP\_Co16, NCI CGAP\_Kid5, NCI CGAP\_Kid12, NCI CGAP\_Kid3, NCI CGAP\_Kid11, NCI CGAP\_Lym2, NCI CGAP\_Br2, NCI CGAP\_Co8, NCI CGAP\_CLL1, NCI CGAP\_Lei2, NCI CGAP\_Brn23, NCI CGAP\_Lu5, NCI CGAP\_Lu24, NCI CGAP\_Lu19, NCI CGAP\_GC4, NCI CGAP\_GC6, NCI CGAP\_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:  
NCI CGAP Kid3 pool 1 LLM 3334-3337, 3682-3683,  
3798-3803 (IMAGE Clones 1322376-1323911, 1456008-1456775,  
1500552-1502855); NCI CGAP Kid5 pool 1 LLM 3338-3342,  
3722-3725, 3776-3778 (IMAGE Clones 1323912-1325831,  
1471368-1472903, 1492104-1493255); NCI CGAP Lu5 pool 1  
LLM 3575-3582, 3851-3854 (IMAGE Clones 1257096-1258631,  
1520904-1522439); NCI CGAP GC4 pool 1 LLM 3164-3167,  
3716-3720, 3733-3735 (IMAGE Clones 1475592-1476743); NCI CGAP Pr22 pool 1  
1469064-1470983, 1475592-1476743; NCI CGAP Pr22 pool 1  
LLM 2457-2459, 3062-3068 (IMAGE Clones  
958608-986759, 1101192-1101959, 1217928-1220615);  
NCI CGAP\_Co10 pool 1 LLM 2644-2653, 2871-2872 (IMAGE  
Clones 1057416-1061255, 114584-1145351). Subtraction  
was performed as previously described (Bonaldo, Lennon &  
Soares (1996): Normalization and Subtraction: Two  
Approaches To Facilitate Gene Discovery. Genome Research  
6, 791-806.  
TAG LIB=NCI CGAP\_Br2  
TAG TISSUE=breast  
TAG SEQ=AAAC"

BASE COUNT 168 a 98 c 98 g 193 t  
ORIGIN

Query Match

57.3%; Score 404.8; DB 10; Length 557;

```

FEATURES
    source
        High quality sequence stop: 523.
        Location/Qualifiers
            1..1150
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5442527"
                /clone_lib="NIH MGC 98"
                /tissue_type="astrocytoma grade IV, cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
                EcoRI; cDNA made by oligo-dT priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGCACGAG(G). Library constructed by Ling Hong
                in the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit

```

RESULT 4	
Bi603219	
LOCUS	
DEFINITION	788 bp mRNA linear EST 07-SEP-2001 603249869F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5301617 5', mRNA sequence.
ACCESSION	Bi603219
VERSION	Bi603219.1 GI:15496158
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo..
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM11763 row: j column: 18 High quality sequence stop: 787.

**FEATURES**  
**source**

Wed May 28 14:44:09 2003

```

/dx xref="taxon:9606"
/clone="IMAGE:5312664"
/clone.lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH108"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      179 a      154 c      181 g      150 t
ORIGIN

```

```

Query Match      35.8%; Score 252.6; DB 13; Length 664;
Best Local Similarity 73.3%; Pred. No. 2.7e-45;
Matches 351; Conservative 0; Mismatches 124; Indels 4; Gaps 2;

QY      6 GAGCTTCGCTCCTGGGAGAGTGGCGGAGAACTGGGGCTCACCGTCATGATGCTCTATC 65
Db      6 GAGCTTCGCTCCTGGGAGAGTGGCGGAGAACTGGGGCTCACCGTCATGATGCTCTATC 65
QY      28 GAGCTTCGCTCCTGGGAGAGTGGCGGAGAACTGGGGCTCACCGTCATGATGCTCTATC 87
Db      28 GAGCTTCGCTCCTGGGAGAGTGGCGGAGAACTGGGGCTCACCGTCATGATGCTCTATC 87
QY      66 AGAAGCAAAATGGCAATTTGCAATTTAACTTTTGAAGAAGCTAGGGGAAACAACTCAA 125
Db      66 AGAAGCAAAATGGCAATTTGCAATTTAACTTTTGAAGAAGCTAGGGGAAACAACTCAA 125
QY      88 AGAAGCAAAATGGCAATTTGCAATTTAACTTTTGAAGAAGCTAGGGGAAACAACTCGAA 147
Db      88 AGAAGCAAAATGGCAATTTGCAATTTAACTTTTGAAGAAGCTAGGGGAAACAACTCGAA 147
QY      126 CAACCTATTTTTC-CCCATGAGCATATCATCAGCTTGGCCATGCTGTTTTCATGGGGGC 184
Db      126 CAACCTATTTTTC-CCCATGAGCATATCATCAGCTTGGCCATGCTGTTTTCATGGGGGC 184
QY      148 GAATGTGTTTTTCTCACCATGAGCATGCTCTGTGCGCCCTGGCCATGGTCTACATGGGGC 207
Db      148 GAATGTGTTTTTCTCACCATGAGCATGCTCTGTGCGCCCTGGCCATGGTCTACATGGGGC 207
QY      185 AAAGGGAACACTGAGCTCAGATGCTCAGGCACTTTTGTGTTAGTAAATCGGAGTGA 244
Db      185 AAAGGGAACACTGAGCTCAGATGCTCAGGCACTTTTGTGTTAGTAAATCGGAGTGA 244
QY      208 AAAGGGAACACTGAGCTCAGATGCTCAGGCACTTTTGTGTTAGTAAATCGGAGTGA 265
Db      208 AAAGGGAACACTGAGCTCAGATGCTCAGGCACTTTTGTGTTAGTAAATCGGAGTGA 265
QY      245 AGATGAGATATTCATCGAGTTCCTGAGGTTTTCAGTCACTTCTGTTGCAATTAACAGAACTGACAC 304
Db      245 AGATGAGATATTCATCGAGTTCCTGAGGTTTTCAGTCACTTCTGTTGCAATTAACAGAACTGACAC 304
QY      266 -GGTGGAGACATCCACAGGGCTTCAGTCTCTTCCAGGAGTGGAGAGCTTGACTTTAT 324
Db      266 -GGTGGAGACATCCACAGGGCTTCAGTCTCTTCCAGGAGTGGAGAGCTTGACTTTAT 324
QY      305 TGAATATGTCTTAGAACTGCCAACGGGCTCTTTGGAGAAAAGTCTTATGATTCTCTAC 364
Db      305 TGAATATGTCTTAGAACTGCCAACGGGCTCTTTGGAGAAAAGTCTTATGATTCTCTAC 364
QY      325 GCAGTACTTGTAGGTTGGCCCAACAGGCTCTTTGGGAAAAGTCTTGTGATTCTCTCTC 384
Db      325 GCAGTACTTGTAGGTTGGCCCAACAGGCTCTTTGGGAAAAGTCTTGTGATTCTCTCTC 384
QY      365 AGTTTTTACAGATTCCTGTCGCAATTTTACCAAGCAACGATAAACAACAGCTAGACTTGT 424
Db      365 AGTTTTTACAGATTCCTGTCGCAATTTTACCAAGCAACGATAAACAACAGCTAGACTTGT 424
QY      385 ATCTTTTAGAGATTCCTGCCAAAATTTTACCAAGCAACGATAAACAACAGCTAGACTTAT 444
Db      385 ATCTTTTAGAGATTCCTGCCAAAATTTTACCAAGCAACGATAAACAACAGCTAGACTTAT 444
QY      425 GAATGATACAGAGAGTCCACAACTGTAACCTCCTGGGTGCTGATGATAAACAAG 483
Db      425 GAATGATACAGAGAGTCCACAACTGTAACCTCCTGGGTGCTGATGATAAACAAG 483
QY      445 CAGCGCGTAGAGAGTCCAGAAAACACATAAACACCTGGGTAGCTGAAAAGACAGAAG 503
Db      445 CAGCGCGTAGAGAGTCCAGAAAACACATAAACACCTGGGTAGCTGAAAAGACAGAAG 503

```

```

RESULT 6
LOCUS   BG749304
DEFINITION 602707953F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4844629 5',
mRNA sequence.
ACCESSION BG749304
VERSION   BG749304.1
KEYWORDS EST.
SOURCE    human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 804)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabs-remail.nih.gov
          Tissue Procurement: ATCC

```

```

BASE COUNT      179 a      154 c      181 g      150 t
ORIGIN

```

```

/dx xref="taxon:9606"
/clone="IMAGE:5312664"
/clone.lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH108"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      220 a      175 c      205 g      187 t
ORIGIN

```

```

Query Match      35.9%; Score 253.2; DB 13; Length 788;
Best Local Similarity 73.3%; Pred. No. 1.9e-45;
Matches 351; Conservative 0; Mismatches 124; Indels 4; Gaps 2;

QY      6 GAGCTTCGCTCCTGGGAGAGTGGCGGAGAACTGGGGCTCACCGTCATGATGCTCTATC 65
Db      6 GAGCTTCGCTCCTGGGAGAGTGGCGGAGAACTGGGGCTCACCGTCATGATGCTCTATC 65
QY      28 GAGCTTCGCTCCTGGGAGAGTGGCGGAGAACTGGGGCTCACCGTCATGATGCTCTATC 87
Db      28 GAGCTTCGCTCCTGGGAGAGTGGCGGAGAACTGGGGCTCACCGTCATGATGCTCTATC 87
QY      66 AGAAGCAAAATGGCAATTTGCAATTTAACTTTTGAAGAAGCTAGGGGAAACAACTCAA 125
Db      66 AGAAGCAAAATGGCAATTTGCAATTTAACTTTTGAAGAAGCTAGGGGAAACAACTCAA 125
QY      88 AGAAGCAAAATGGCAATTTGCAATTTAACTTTTGAAGAAGCTAGGGGAAACAACTCGAA 147
Db      88 AGAAGCAAAATGGCAATTTGCAATTTAACTTTTGAAGAAGCTAGGGGAAACAACTCGAA 147
QY      126 CAACCTATTTTTC-CCCATGAGCATATCATCAGCTTGGCCATGCTGTTTTCATGGGGGC 184
Db      126 CAACCTATTTTTC-CCCATGAGCATATCATCAGCTTGGCCATGCTGTTTTCATGGGGGC 184
QY      148 GAATGTGTTTTTCTCACCATGAGCATGCTCTGTGCGCCCTGGCCATGGTCTACATGGGGC 207
Db      148 GAATGTGTTTTTCTCACCATGAGCATGCTCTGTGCGCCCTGGCCATGGTCTACATGGGGC 207
QY      185 AAAGGGAACACTGAGCTCAGATGCTCAGGCACTTTTGTGTTAGTAAATCGGAGTGA 244
Db      185 AAAGGGAACACTGAGCTCAGATGCTCAGGCACTTTTGTGTTAGTAAATCGGAGTGA 244
QY      208 AAAGGGAACACTGAGCTCAGATGCTCAGGCACTTTTGTGTTAGTAAATCGGAGTGA 265
Db      208 AAAGGGAACACTGAGCTCAGATGCTCAGGCACTTTTGTGTTAGTAAATCGGAGTGA 265
QY      245 AGATGAGATATTCATCGAGTTCCTGAGGTTTTCAGTCACTTCTGTTGCAATTAACAGAACTGACAC 304
Db      245 AGATGAGATATTCATCGAGTTCCTGAGGTTTTCAGTCACTTCTGTTGCAATTAACAGAACTGACAC 304
QY      266 -GGTGGAGACATCCACAGGGCTTCAGTCTCTTCCAGGAGTGGAGAGCTTGACTTTAT 324
Db      266 -GGTGGAGACATCCACAGGGCTTCAGTCTCTTCCAGGAGTGGAGAGCTTGACTTTAT 324
QY      305 TGAATATGTCTTAGAACTGCCAACGGGCTCTTTGGAGAAAAGTCTTATGATTCTCTAC 364
Db      305 TGAATATGTCTTAGAACTGCCAACGGGCTCTTTGGAGAAAAGTCTTATGATTCTCTAC 364
QY      325 GCAGTACTTGTAGGTTGGCCCAACAGGCTCTTTGGGAAAAGTCTTGTGATTCTCTCTC 384
Db      325 GCAGTACTTGTAGGTTGGCCCAACAGGCTCTTTGGGAAAAGTCTTGTGATTCTCTCTC 384
QY      365 AGTTTTTACAGATTCCTGTCGCAATTTTACCAAGCAACGATAAACAACAGCTAGACTTGT 424
Db      365 AGTTTTTACAGATTCCTGTCGCAATTTTACCAAGCAACGATAAACAACAGCTAGACTTGT 424
QY      385 ATCTTTTAGAGATTCCTGCCAAAATTTTACCAAGCAACGATAAACAACAGCTAGACTTAT 444
Db      385 ATCTTTTAGAGATTCCTGCCAAAATTTTACCAAGCAACGATAAACAACAGCTAGACTTAT 444
QY      425 GAATGATACAGAGAGTCCACAACTGTAACCTCCTGGGTGCTGATGATAAACAAG 483
Db      425 GAATGATACAGAGAGTCCACAACTGTAACCTCCTGGGTGCTGATGATAAACAAG 483
QY      445 CAGCGCGTAGAGAGTCCAGAAAACACATAAACACCTGGGTAGCTGAAAAGACAGAAG 503
Db      445 CAGCGCGTAGAGAGTCCAGAAAACACATAAACACCTGGGTAGCTGAAAAGACAGAAG 503

```

```

RESULT 5
LOCUS   BI669663
DEFINITION 603293204F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312664 5',
mRNA sequence.
ACCESSION BI669663
VERSION   BI669663.1
KEYWORDS EST.
SOURCE    human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 664)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabs-remail.nih.gov
          Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
          CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
          Toshiyuki and Piero Carninci (RIKEN)
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L1AM11792 row: g column: 01
          High quality sequence stop: 659.
          Location/Qualifiers
            1..664
              /organism="Homo sapiens"

```

```

BASE COUNT      220 a      175 c      205 g      187 t
ORIGIN

```

CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM1682 row: i column: 14  
 High quality sequence stop: 683.

FEATURES  
 source  
 1. .804  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4844629"  
 /clone\_lib="NIH\_MGC\_43"  
 /tissue\_type="normal pigmented retinal epithelium"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library." |

BASE COUNT 229 a 187 c 205 g 183 t

Query Match 35.8%; Score 252.6; DB 12; Length 804;  
 Best Local Similarity 73.3%; Pred. No. 2.5e-45;  
 Matches 351; Conservative 0; Mismatches 124; Indels 4; Gaps 2;

QY 6 GAGCTTCGCTCCTGGGAGCTGCGGAGAACTGGGCTACCGTCATGATGCTCTATC 65  
 DB 10 GAGCTTCGCGAGGAGAGACGGCGCGCGGCTGCGCATCATGATGCTTCGC 69

QY 66 AGAAGCAAAATGGCAATTTGATTAACCTTTTGAAGAACTAGGGGAAACAACTCAA 125  
 DB 70 AGAAGCAAAATGGCACTTTGCTTAAACCTTTTGAAGAACTAGGGTAAAGACAACTCGAA 129

QY 126 CAATCTATTTTTTC-CCCATGAGCATATCATCAGCCTTGGCCATGTTTTCATGGGGC 184  
 DB 130 GAATGTGTTTTCTCACCCATGAGCATGTCTGTGCGCTGCGCATGCTTACATGGGGC 189

QY 185 AAAGGAAACACTCAGCTCAGATGTCTCAGGACATTTGTTTGTAGTAAATCGGAGTGA 244  
 DB 190 AAAGGAAACACTCAGCTCAGATGTCTCAGGACATTTTCTTCAATAAAGTGGCGT-- 247

QY 245 AGATGAGATATTTCATCGAGGTTTTTCAGTCACATCTTGTGTCATTAACAGAACTGCAC 304  
 DB 248 -GGTGGAGACATCCACCAGGCTTCCAGTCTCTTCCAGCAAGTGAACAGACTGGCAC 306

QY 305 TGAATATGCTTTAGAACTCCAAAGGCTCTTTGGAGAAAGTCTTATGATTTCTCTCAC 364  
 DB 307 GCAGTACTTCTTAGGATGGCCAAAGGCTCTTTGGGAAAGTCTTGTGATTTCTCTC 366

QY 365 AGGTTTTACAGATTTCTGTGGCAAAATTTACCAAGCAACGATAAACAGCTAGACTTGT 424  
 DB 367 ATCTTTTAGAGATTTCTGCCAAAATTTCTACCAAGCAGAGATGGAGAGCTTGACTTTAT 426

QY 425 GAATGATACAGAGAGTCCCAACAGCTGTAAACTCTCTGGTTGCTGATAAAACTAAAG 483  
 DB 427 CAGCGCGGTAGAGAGTCCAGAAACACATAAACCTCTGGTAGCTGAAAGACAGAAAG 485

RESULT 7  
 LOCUS B0681973  
 DEFINITION AGENCOURT\_9212433 NIH\_MGC\_112 Homo sapiens CDNA clone IMAGE:6260725  
 5', mRNA sequence.  
 ACCESSION B0681973  
 VERSION B0681973.1 GI:21794652  
 KEYWORDS EST.  
 SOURCE human.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 905)

## AUTHORS

NIH-MGC http://mgi.nci.nih.gov/

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2420 row: e column: 14

High quality sequence stop: 649.

## FEATURES

source

1. .905

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:6260725"

/clone\_lib="NIH\_MGC\_112"

/tissue\_type="melanotic melanoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCACGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH\_MGC Library."

BASE COUNT 261 a 196 c 246 g 201 t 1 others

## Query Match

Best Local Similarity 35.8%; Score 252.6; DB 14; Length 905;

Matches 351; Conservative 0; Mismatches 124; Indels 4; Gaps 2;

QY

6 GAGCTTCGCTCCTGGGAGCTGCGGAGAACTGGGCTACCGTCATGATGCTCTATC 65

DB

45 GAGCTTCGCGAGGAGAGAACGGCGCGCGGCTCGCCATCATGATGCTCTGC 104

QY

66 AGAAGCAAAATGGCAATTTGATTAACCTTTTGAAGAACTAGGGGAAACAACTCAA 125

DB

105 AGAAGCAAAATGGCACTTTGCTTAAACCTTTTGAAGAACTAGGGTAAAGACAACTCGAA 164

QY

126 CAATCTATTTTTTC-CCCATGAGCATATCATCAGCCTTGGCCATGTTTTCATGGGGC 184

DB

165 GAATGTGTTTTCTCACCCATGAGCATGTCTGTGCGCTGCGCATGTTTACATGGGGC 224

QY

185 AAAGGAAACACTCAGCTCAGATGTCTCAGGACATTTGTTTGTAGTAAATCGGAGTGA 244

DB

225 AAAGGAAACACTCAGCTCAGATGTCTCAGGACATTTTCTTCAATAAAGTGGCGT-- 282

QY

245 AGATGAGATATTTCATCGAGGTTTTTCAGTCACATCTTGTGTCATTAACAGAACTGCAC 304

DB

283 -GGTGGAGACATCCACCAGGCTTCCAGTCTCTTCTCAGCAAGTGAACAGACTGGCAC 341

QY

305 TGAATATGCTTTAGAACTCCAAAGGCTCTTTGGAGAAAGTCTTATGATTTCTCTCAC 364

DB

342 GCAGTACTTCTTAGGATGGCCAAAGGCTTTTGGGAAAGTCTTGTGATTTCTCTC 401

QY

365 AGGTTTTACAGATTTCTGTGGCAAAATTTTACCAAGCAACGATAAACAGCTAGACTTGT 424

DB

402 ATCTTTTAGAGATTTCTGCCAAAATTTCTACCAAGCAGAGATGGAGAGCTTGACTTTAT 461

QY

425 GAATGATACAGAGAGTCCCAACAGCTGTAAACTCTCTGGTTGCTGATAAAACTAAAG 483

DB

462 CAGCGCGGTAGAGAGTCCAGAAACACATAAACCTCTGGTAGCTGAAAGACAGAAAG 520

372 ACAGATTCTGTGGCAAAATTTACCAAGCAACGATAAAACAGCTAGACTTTGTGAATGAT 431  
 379 AAAGATTCTCTGCCCAAAATTTACCAAGCGAGATGGAAGCTGGACTTTGTCAACACG 438  
 432 ACAGAGAAAGTCCACACACAGCTGTAAACTCTCTGGTGTCTGTATATAAACTAAAG 483  
 439 ACAGAGGAGTCCCGGAAACACATAAACACCTGGGTAGCCGAGAGACAGAAG 490

RESULT 9  
 BE543971 536 bp mRNA linear EST 09-AUG-2000  
 LOCUS 601070485F1 NIH\_MGC\_12 Homo sapiens cDNA clone IMAGE:3456500 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE543971  
 VERSION BE543971.1 GI:9772616  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 536)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: Incyte Genomics, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LAM8444 row: j column: 21  
 High quality sequence stop: 534.  
 Location/Qualifiers  
 1..536  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="IMAGE:3456500"  
 /clone\_lib="NIH MGC 12"  
 /tissue\_type="cervical carcinoma cell line"  
 /lab\_host="DH10B"  
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site: 1; NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.4 kb. Library prepared by Life  
 Technologies."  
 BASE COUNT 132 a 143 c 135 g 126 t  
 ORIGIN  
 Query Match 35.6%; Score 251.6; DB 10; Length 536;  
 Best Local Similarity 73.8%; Pred. No. 4.7e-45;  
 Matches 347; Conservative 0; Mismatches 119; Indels 4; Gaps 2;  
 15 TCCTGGGAGCTCGCGGAGAACTGGGCTCCCGCTCATGGATGCTCTATCAGAACAAA 74  
 61 TTCTGTCTGCTCGCTCCCGCTCTGGAGTCTGGATCATGGATGTTCTCGAGAACAAA 120  
 75 TGGCACATTTGGCAATTAACCTTTTGAAGAAAGCTAGGGGAAACAACTCAACAACTTAT 134  
 121 TGGACCTTTGGCTTAAACCTTTTGAAGAAAGCTAGGGGAAACAACTCAAGAAATGTGT 180  
 135 TTTTTC-CCCATGAGCATATCATCAGCTTGGCCATGGTTTCATCGGGCGAAGGAAA 193  
 181 TTTTCTACCCATGAGCATGTCTGTGCTGGCCATGGTCTATCGGGCGAAGGAAA 240  
 194 CACTGAGCTCAGATGTCTCAGGACCTTTGTTTGTAGTAAATCGAGGTGAAGATGAGA 253  
 241 CACCGCTGACAGATGCGCCAGATATCTTCTTTCAATAAAGTGGCGGT---GGTGA 297  
 254 TATTTCAGAGGTTTTCAGTCACTTCTGTGTTGCAATTAACAGAACTGCACTGAATATGT 313

372 ACAGATTCTGTGGCAAAATTTACCAAGCAACGATAAAACAGCTAGACTTTGTGAATGAT 431  
 379 AAAGATTCTCTGCCCAAAATTTACCAAGCGAGATGGAAGCTGGACTTTGTCAACACG 438  
 432 ACAGAGAAAGTCCACACACAGCTGTAAACTCTCTGGTGTCTGTATATAAACTAAAG 483  
 439 ACAGAGGAGTCCCGGAAACACATAAACACCTGGGTAGCCGAGAGACAGAAG 490

RESULT 8  
 BG894979 595 bp mRNA linear EST 05-JUN-2001  
 LOCUS 355775 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION BG894979  
 VERSION BG894979.1 GI:14305220  
 KEYWORDS EST.  
 SOURCE pig.  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 REFERENCE 1 (bases 1 to 595)  
 AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,  
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.  
 and Keele,J.W.  
 TITLE Design and use of two pooled tissue normalized cDNA libraries for  
 EST discovery in swine  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called and alt-trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -minscore 18  
 and -minmatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCAGTCAAGCAGC  
 Plate: 120 row: P column: 8  
 Seq primer: ATTAGGTGACATATAG.  
 Location/Qualifiers  
 1..595  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /clone\_lib="MARC 1P1G"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site: 1; NotI; Site: 2: SalI;  
 Library made from pooled tissue from day 11, 13, 15, 20,  
 and 30 embryos."  
 BASE COUNT 157 a 173 c 136 g 129 t  
 ORIGIN  
 Query Match 35.7%; Score 252; DB 12; Length 595;  
 Best Local Similarity 72.2%; Pred. No. 3.7e-45;  
 Matches 341; Conservative 0; Mismatches 130; Indels 1; Gaps 1;  
 12 CGCTCTGGGAGCTGCGGAGAACTGGGCTCACCGTCACTGATGCTCTATCAGAAAGC 71  
 20 CGCCCGCCCGCAGCTCGGCTCGGCTCGGACTTGCCTCATGATGCTCTGTCGGAGGC 79  
 72 AATGGCAGATTGTCATTAACCTTTTGAAGAAAGCTAGGGGAAACAACTCAACAACTT 131  
 80 AATGGCAGATTGTCATTAACCTTTTGAAGAAAGCTAGGGGAAACAACTCAACAACTT 138  
 132 ATTTTTCCTCCATGAGCATATCATCAGCTTGGCCATGGTTTCATCGGGCGAAGGGA 191  
 139 TTTTCTACCCATGAGCTCTCTCCAGCTGGCCATGGTCTTTATGGGGCGAGCAGGC 198  
 192 AACCTGAGCTCAGATGTCTCAGGACCTTTGTTTGTAGTAAATCGAGGTGAAGATGGA 251  
 199 AACACCGCAACTCAGATGTCAGGACCTTTCTTAAGCAAAAGCAGCAGCAAGGTGGA 258  
 252 GATATTTCATCGAGGTTTTCAGTCACTTTCTGTTGCAATTAACAGAACTGCACTGAATAT 311  
 259 GATGTCACACAGGTTTTCAGAACTTCTCAACCGAAGTCAACAGGACTGGCACAGTAC 318  
 312 GTGCTTAGACTGCCAACGGCTTTTGGAGAAAGTCTTATGATTCTCTCAGAGGTTT 371  
 319 TTGCTCAGAAACCGCCAAACAGGCTCTTCGGAGAGAGTCTTACGGTTTCTCTCCATCTTC 378

```

Db      298 CATCCACAGGGCTCTCCAGTCTCTCTCACCAGAGTGAACAGACTGGCACCAGTACTT 357
Qy      314 GCTTAGAAGTGCACCAAGGGCTCTTTGGAGAAAGTCTTATGATTTCTCCTCAGAGTTTAC 373
Db      358 GCTTAGAGTGGCCACACAGCTCTTTGGGAAAGTCTTTGATTTCTCTCATCTTTTAC 417
Qy      374 AGATTCTGTGGCAAAATTTACCAAGCAACGATAAACAGCTAGACTTTGTGAATGATAC 433
Db      418 AGATTCTGTGGCAAAATTTACCAAGCAACGATGAGGAGCTTGATTTATCAGGCGGT 477
Qy      434 AGAGAAGTCCCAACACAGCTGTAACCTCTGGGTTGCTGATATAAACTAAAG 483
Db      478 AGAGAAGTCCCAACACACATATAACACCTGGGTAGTGAAGAGCAGAAG 527

RESULT 10
LOCUS      BI222640                818 bp      mRNA      linear      EST 11-JUL-2001
DEFINITION 602940135F1 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:5103344 5',
mRNA sequence.
ACCESSION  BI222640
VERSION     BI222640.1 GI:14676084
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE     NIH-MGC http://mgc.nci.nih.gov/
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11249 row: e column: 09
High quality sequence stop: 764.
FEATURES   Location/Qualifiers
            1..818
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5103344"
            /clone_lib="NIH_MGC_12"
            /tissue_type="cervical carcinoma cell line"
            /lab_host="DH10B"
            /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.4 kb. Library prepared by Life
            Technologies."
BASE COUNT 191 a 224 c 221 g 182 t
ORIGIN
Query Match 35.6%; Score 251.6; DB 13; Length 818;
Best Local Similarity 73.8%; Pred. No. 4.2e-45;
Matches 347; Conservative 0; Mismatches 119; Indels 4; Gaps 2;
Qy      15 TCCTGGCAGCTGCGGGAGAACTGGGGCTACCGTCACTGGATGCTCTATCAGAGCAAA 74
Db      158 TTCTGCTCGCTCCCCGCTCTGGAGTCTGCCATCATGGATGTTCTGCAGAGCAAA 217
Qy      75 TGSCACATTGGCATTAACCTTTTGAAGAAAGCTAGGGGAAACAACTCAAACTTATT 134
Db      218 TGSCACCTTTGGCTTAAACCTTTTGAAGAAAGCTGGGTAAGAACTCGAAGATGTGT 277
Qy      135 TTTTTC-CCCATGAGCATATCATCAAGCTTTGGCCATGGTTTTCATGGGGCAAGGAAA 193
Db      278 TTTCTCACCCATGAGCATGTCTGTGCTGCTGCGCCATGGTCTACATGGGGCAAGGAAA 337

```

```

Qy      194 CACTGCAGCTCAGATGTCTCAGGCACCTTTGTTTAAATCGGAGGTGAAGATGAGA 253
Db      338 CACCGTGCACAGATGGCCAGATACTTTCTTTCAATAAAGTGCGGT---GGTGAGA 394
Qy      254 TATTTCATCAGAGTTTTCAGTCACTTCTTGTGCAATTACAGAACTGACACTGAATATGT 313
Db      395 CATCCACAGGGCTTCCAGTCTCTTCTCACCCAGAGTGAACAAGACTGGCAGCAGTACTT 454
Qy      314 GCTTAGAAGTGCACCAAGGGCTCTTTGGAGAAAGTCTTATGATTTCTCCTCAGAGTTTAC 373
Db      455 GCTTAGAGTGGCCACACAGGCTCTTTGGGAAAGTCTTGTGATTTCTCTCATCTTTTAC 514
Qy      374 AGATTCTGTGGCAAAATTTACCAAGCAACGATAAACAGCTAGACTTTGTGAATGATAC 433
Db      515 AGATTCTGTGGCAAAATTTACCAAGCAACGATGAGGAGCTTGACTTTATCAGCCTCGT 574
Qy      434 AGAGAAGTCCCAACACAGCTGTAACCTCTGGGTTGCTGATATAAACTAAAG 483
Db      575 AGAGAAGTCCCAACACATATAACACCTGGGTAGTGAAGAGCAGAAG 624

RESULT 11
LOCUS      BG108578                875 bp      mRNA      linear      EST 30-JAN-2001
DEFINITION 602278558F1 NIH_MGC_86 Homo sapiens CDNA clone IMAGE:4366047 5',
mRNA sequence.
ACCESSION  BG108578
VERSION     BG108578.1 GI:12602424
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE     NIH-MGC http://mgc.nci.nih.gov/
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10016 row: d column: 16
High quality sequence stop: 704.
FEATURES   Location/Qualifiers
            1..875
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4366047"
            /clone_lib="NIH_MGC_86"
            /tissue_type="osteosarcoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: bone; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
            Average insert size 1.533 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH_MGC Library."
BASE COUNT 224 a 216 c 225 g 210 t
ORIGIN
Query Match 35.6%; Score 251.6; DB 12; Length 875;
Best Local Similarity 73.8%; Pred. No. 4.1e-45;
Matches 347; Conservative 0; Mismatches 119; Indels 4; Gaps 2;
Qy      15 TCCTGGCAGCTGCGGGAGAACTGGGGCTCAGGTCATGATGCTCTATCAGAGCAAA 74
Db      76 TTCTGCTCGCTCCCCGCTCTGGAGTCTGCCATCATGGATGTTCTCGCAGAGCAAA 135

```

QY 75 TGGCACAATTGGCATTAAACCTTTTGAAGAGCTAGGGGAAACAACTCAACAACTTATT 134  
 Db 136 TGGCACCCTTTGCTTAAACCTTTTGAAGAGCTAGGGGAAACAACTCAACAACTTATT 195  
 QY 135 TTTTTC-CCCATGAGCATATCATCAGCTTGGCCATGGTTTTCATGGGGCAAGGAAA 193  
 Db 196 TTTTCTCACCCTATGAGCATGTCTGTGCTTGGCCATGGTTTTCATGGGGCAAGGAAA 255  
 QY 194 CACTGCAGCTCAGATGTCTCAGGCACTTTTGTAGTAAATCGGAGTGAAGTGGAGA 253  
 Db 256 CACCGTGCACAGATGGCCAGATACTTTCTTCAATAAAAGTGGCGGT---GGTGGAGA 312  
 QY 254 TATTTCATCGAGGTTTTCAGTCACTTCTTGTGCAATTAACAGACTGACACTGAATATGT 313  
 Db 313 CATCCACCGGCTTCCAGTCTCTTCTCACCAGAGTGAACAGACTGGCAGCAGTACTT 372  
 QY 314 GCTTAGACTGCCAAGCGGCTTTTGGAGAAAGTCTTATGATTTTCTCAGAGTTTATAC 373  
 Db 373 GCTTAGAGTGGCCAAACAGGCTTTTGGGAAAGTCTTGTGATTTCTCTCATCTTTTAG 432  
 QY 374 AGATTCCTGTGGCAAAATTTTACCAAGCAACGATATAACAGCTAGACTTTGTGAATATAC 433  
 Db 433 AGATTCCTGTGGCAAAATTTTACCAAGCAGAGATGGAGGAGCTTGAATTTATCAGCGCGT 492  
 QY 434 AGAGAAGTCCAGAAACACATCAACCTCTGGTGTGCTGATAAACTAAAG 483  
 Db 493 AGAGAAGTCCAGAAACACATCAACCTCTGGTGTGCTGATAAACTAAAG 542

RESULT 12  
 BM461676 880 bp mRNA linear EST 05-FEB-2002  
 LOCUS AGENCOURT 6418371 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5533721  
 5', mRNA sequence.

ACCESSION BM461676  
 VERSION BM461676.1 GI:18510716  
 KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 880)  
 NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12219 row: a column: 18

High quality sequence stop: 659.

Location/Qualifiers

1..880

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5533721"

/clone.lib="NIH MGC 71"

/tissue\_type="leiomyosarcoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: Not I; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 2.1 kb.

BASE COUNT 239 a 206 c 226 g 209 t

ORIGIN

Query Match 35.6%; Score 251.6; DB 13; Length 880;

Best Local Similarity 73.8%; Pred. No. 4.1e-45;

Matches 347; Conservative 0; Mismatches 119; Indels 4; Gaps 2;

QY 15 TCCTGGCAGCTGCGGGAGAACTGGGGCTCACCGTCATGATGCTCTATCAGAAGCAA 74  
 Db 74 TTTCTGCTCGTCTGCCCTCTGGAGTCTGCCATCATGATGTTCTCGCAGAAGCAA 133  
 QY 75 TGGCACAATTTGCATTAAACCTTTTGAAGAGCTAGGGGAAACAACTCAACAACTTATT 134  
 Db 134 TGGCACCCTTTGCTTAAACCTTTTGAAGAGCTAGGGGAAACAACTCAACAACTTATT 193  
 QY 135 TTTTTC-CCCATGAGCATATCATCAGCTTGGCCATGGTTTTCATGGGGCAAGGAAA 193  
 Db 194 TTTTCTCACCCTATGAGCATGTCTGTGCTTGGCCATGGTTTTCATGGGGCAAGGAAA 253  
 QY 194 CACTGCAGCTCAGATGTCTCAGGCACTTTTGTAGTAAATCGGAGTGAAGTGGAGA 253  
 Db 254 CACCGTGCACAGATGGCCAGATACTTTCTTCAATAAAAGTGGCGGT---GGTGGAGA 310  
 QY 254 TATTTCATCGAGGTTTTCAGTCACTTCTTGTGCAATTAACAGACTGACACTGAATATGT 313  
 Db 311 CATCCACCGGCTTCCAGTCTCTTCTCACCAGAGTGAACAGACTGGCAGCAGTACTT 370  
 QY 314 GCTTAGACTGCCAAGCGGCTTTTGGAGAAAGTCTTATGATTTTCTCAGAGTTTATAC 373  
 Db 371 GCTTAGAGTGGCCAAACAGGCTTTTGGGAAAGTCTTGTGATTTCTCTCATCTTTTAG 430  
 QY 374 AGATTCCTGTGGCAAAATTTTACCAAGCAACGATATAACAGCTAGACTTTGTGAATATAC 433  
 Db 431 AGATTCCTGTGGCAAAATTTTACCAAGCAGAGATGGAGGAGCTTGAATTTATCAGCGCGT 490  
 QY 434 AGAGAAGTCCAGAAACACATCAACCTCTGGTGTGCTGATAAACTAAAG 483  
 Db 491 AGAGAAGTCCAGAAACACATCAACCTCTGGTGTGCTGATAAACTAAAG 540

RESULT 13

BM461676

LOCUS AGENCOURT 8744336

DEFINITION IMAGE:6205988 5', mRNA sequence.

ACCESSION BM461676

VERSION BM461676.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 927)

NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Dr. James R. Lupski

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13629 row: 1 column: 21

High quality sequence stop: 729.

Location/Qualifiers

1..927

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:6205988"

/clone.lib="Lupski\_sciatic\_nerve"

/sex="male"

/tissue\_type="sciatic nerve"

/dev\_stage="adult, 70 yr"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:

NotI; Site 2: SalI; cDNA made by oligo-dT priming.



Directionally cloned using the following adaptors:  
 5'-TGACCCGCGTCGCG-3' and  
 5'-GATAGTTCTAGTCGCGCGCCCT(15)-3'. Size selected >  
 1 kb for average insert length 1.87 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupsaki, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life  
 Technologies."

BASE COUNT 253 a 208 c 244 g 220 t 2 others  
 ORIGIN

Query Match 35.6%; Score 251.6; DB 14; Length 927;  
 Best Local Similarity 73.8%; Pred. No. 4e-45;  
 Matches 347; Conservative 0; Mismatches 119; Indels 4; Gaps 2;

QY 15 TCCTGGCAGCTGCGCGAGAACTGGGGCTCACCGTCATGGATGCTCTATCAGAAGCAAA 74  
 |||||  
 Db 58 TTCTGCTGGCTCGCTCCCGCTCTGGAGTCTGCCATCATGATGTTCTCGCAGAGCAAA 117  
 |||||  
 QY 75 TGGCAGATTTCGATTAAACCTTTTGAAGAGCTAGGGGAAACAACTCAAACTTATT 134  
 |||||  
 Db 118 TGGCAGCTTTGGCTTTAAACCTTTTGAAGAGCTAGGGGAAACAACTCGAAGATGTGT 177  
 |||||  
 QY 135 TTTTTC-CCCATGAGCATATCATAGCCTTGCGCATGTTTTCATGGGGCAAGGAAA 193  
 |||||  
 Db 178 TTTCTACCATGAGCATGCTCTGTGCCCTGGCCATGCTTACATGGGGGCAAGGAAA 237  
 |||||  
 QY 194 CACTGCAGCTCAGATGCTCAGGCACCTTTGTTTGTAGTAAATCGAGGTGAAGATGAGA 253  
 |||||  
 Db 238 CACCGCTGCACAGATGGCCAGATATCTTTTCAATAAAGTGGCGT---GGTGGAGA 294  
 |||||  
 QY 254 TATTATCATGAGTTCAGTCACCTCTTGTGCAATTAACAGACTGACACTGAATATGT 313  
 |||||  
 Db 295 CATCCACAGGCTTCCAGTCTCTCTCACCGAAGTGAACAGACTGCGCAGCAGTACTT 354  
 |||||  
 QY 314 GCTTAGAATGCCAACCGGCTCTTTGGAGAAAGCTTATGATTTCTCACAGTCTTTAC 373  
 |||||  
 Db 355 GCTTAGATGGCCAACAGGCTCTTTGGGAAAGCTTGTGATTTCTCTCATCTTTTAG 414  
 |||||  
 QY 374 AGATTCCTGTGGCAATTTCTACCAAGCAACGATAAAGCAGTACGTTTGTGAATGATAC 433  
 |||||  
 Db 415 AGATTCTGTCCAAAATTTCTACCAAGCAGAGATGGAGGAGCTTGACTTTATCAGCGCGT 474  
 |||||  
 QY 434 AGAGAAGTCCACACACAGTGTAAACTCTCGGTTGCTGATAAACTAAAG 483  
 |||||  
 Db 475 AGAGAAGTCCAGAAACACATAAACACCTGGGTAGCTGAAAGACAGAA 524  
 |||||

RESULT 14  
 Bi822888  
 LOCUS 603040326F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5181246 5',  
 DEFINITION mRNA sequence.  
 ACCESSION Bi822888  
 VERSION Bi822888.1 GI:15934438  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC <http://mhc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs@femail.nih.gov](mailto:cgabs@femail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLAM11452 row: c column: 07  
 High quality sequence stop: 797.

FEATURES  
 source

Location/Qualifiers  
 1..957  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5181246"  
 /clone\_lib="NIH\_MGC\_115"  
 /lab\_host="DH10B"  
 /note="Organ: pooled brain, lung, testis; Vector:  
 pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
 source anonymous pool of 6 male brains, age range 23-27; 1  
 male lung, age 27; and 1 male testis, age 69. Library is  
 oligo-dT primed and directionally cloned (EcoRV site is  
 destroyed upon cloning). Average insert size 1.8 kb,  
 insert size range 1-3 kb. Library is normalized and  
 enriched for full-length clones and was constructed by C.  
 Gruber (Invitrogen). Research Genetics tracking code  
 021. Note: this is a NIH MGC Library."  
 BASE COUNT 223 a 256 c 260 g 218 t  
 ORIGIN

Query Match 35.6%; Score 251.6; DB 13; Length 957;  
 Best Local Similarity 73.8%; Pred. No. 4e-45;  
 Matches 347; Conservative 0; Mismatches 119; Indels 4; Gaps 2;

QY 15 TCCTGGCAGCTGCGCGAGAACTGGGGCTCACCGTCATGGATGCTCTATCAGAAGCAAA 74  
 |||||  
 Db 203 TTTCTGCTCGCTCGCTCCCGCTCTGGAGTCTGCCATCATGATGTTTCTCGCAGAGCAAA 262  
 |||||  
 QY 75 TGGCAGATTTCGATTAAACCTTTTGAAGAGCTAGGGGAAACAACTCAAACTTATT 134  
 |||||  
 Db 263 TGGCAGCTTTGGCTTTAAACCTTTTGAAGAGCTAGGGGAAACAACTCGAAGATGTGT 322  
 |||||  
 QY 135 TTTTTC-CCCATGAGCATATCATAGCCTTGCGCATGTTTTCATGGGGCAAGGAAA 193  
 |||||  
 Db 323 TTTTTCACCCATGAGCATGCTCTGTGCCCTGGCCATGTTTCTACATGGGGCAAGGAAA 382  
 |||||  
 QY 194 CACTGCAGCTCAGATGCTCAGGCACCTTTGTTTGTAGTAAATCGAGGTGAAGATGAGA 253  
 |||||  
 Db 383 CACCGCTGCACAGATGGCCAGATATCTTTTCAATAAAGTGGCGGT---GGTGGAGA 439  
 |||||  
 QY 254 TATTATCATGAGTTCAGTCACTCTTTTGTGCAATTAACAGAACTGACACTGAATATGT 313  
 |||||  
 Db 440 CATCCACAGGCTTCCAGTCTCTTCTCACCGAAGTGAACAGACTGCGCAGCAGTACTT 499  
 |||||  
 QY 314 GCTTAGAATGCCAACCGGCTCTTTGGAGAAAGCTTATGATTTCTCTCACAGTCTTTAC 373  
 |||||  
 Db 500 GCTTAGGATGGCCAACAGGCTCTTTGGGAAAGCTTTGTGATTTCTCTCATCTTTTAG 559  
 |||||  
 QY 374 AGATTCTGTGGCAATTTCTACCAAGCAACGATAAACAGCTAGACTTTGTGAATGATAC 433  
 |||||  
 Db 560 AGATTCTGTGGCAATTTCTACCAAGCAGAGATGGAGGAGCTTGACTTTATCAGCGCGT 619  
 |||||  
 QY 434 AGAGAAGTCCACACACAGCTGTAAACTCTCGGTTGCTGATAAACTAAAG 483  
 |||||  
 Db 620 AGAGAAGTCCAGAAACACATAAACACCTGGGTAGCTGAAAGACAGAG 669  
 |||||

RESULT 15  
 Bi822888  
 LOCUS 603040326F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5761538  
 DEFINITION 5', mRNA sequence.  
 ACCESSION Bi822888  
 VERSION Bi822888.1 GI:19901431  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC <http://mhc.nci.nih.gov/>.

Wed May 28 14:44:09 2003

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LAM2810 row: f column: 03  
High quality sequence stop: 666.  
Location/Qualifiers  
1. .991  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5761538"  
/clone\_lib="NIH\_MGC\_122"  
/lab\_host="DH10B"  
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;  
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source  
anonymous pool of 24 week female lung, 16 week female  
spleen, and 20-22 week male spleens. Library is oligo-dt  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.4 kb, insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 026. Note:  
this is a NIH MGC Library."  
275 a 230 c 255 g 231 t

BASE COUNT  
ORIGIN  
Query Match 35.6%; Score 251.6; DB 14; Length 991;  
Best Local Similarity 73.8%; Pred. No. 4e-45;  
Matches 347; Conservative 0; Mismatches 119; Indels 4; Gaps 2;  
Qy 15 TCCTGGGCGAGCTGGCGGAGAACCTGGGGCTCACCGTCATGATGCTCTATCAGAAGCAAA 74  
Db 80 TTCTGCTCGCTCGCTCCCGCTCTGGAGTCGCGCATCATGATGTTCTCGCAGAGCAAA 139  
Qy 75 TGGCACATTTTCATTAAACCTTTTGAAGAGCTAGGGGAAACAACTCAACAACTTATT 134  
Db 140 TGGACCTTTGCTTTAAACCTTTTGAAGAGCTAGGGGAAACAACTCAACAACTTATT 199  
Qy 135 TTTTTC-CCAGAGCATATCATCAGCTTGGCCATGTTTTCATGGGGCAAGGCAAA 193  
Db 200 TTTTCTCACCATGAGCATGCTCTGTCGCCCTGGCCATGTTTTCATGGGGCAAGGCAAA 259  
Qy 194 CACTGCAGCTCAGATGCTCAGGCACCTTTGTTTGTAGTAAATCGAGGTGAAGTGGAGA 253  
Db 260 CACCGCTGACAGATGGCCAGATACCTTTCTTCAATAAAGTGCGGT---GGTGGAGA 316  
Qy 254 TATTCATCGAGGTTTTCAGTCACTTTGTTGGCAATTAACAGAACTGACACTGAATATGT 313  
Db 317 CATCCACAGGCTTCCAGTCTCTCTCAGGAAGTGAACAGACTGGCAGCGACTT 376  
Qy 314 GCTTAGAAGTCCACAGGCTTTTGGAGAAAGTCTTATGATTTCTTCACAGGTTTAC 373  
Db 377 GCTTAGAAGTCCACAGGCTTTTGGGAAAGTCTTGTGATTTCTCTCTCTTTAG 436  
Qy 374 AGATTTCCTGTGGCAATTTACCAAGCAACGATTAACAAAGCTAGACTTTGTGAATGATAC 433  
Db 437 AGATTTCCTGCCAAAATTTTACCAAGCAGAGATGGAGGCTTGACTTTATCAGCGCGT 496  
Qy 434 AGAGAAGTCCACAAACAGCTGTAACCTCTGGGTTCTCTGATTAATAAG 483  
Db 497 AGAGAAGTCCAGAAAACACATAAACACCTGGGTAGTGTGAAAGACAGAAAG 546

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 10:19:06 ; Search time 76 seconds  
(without alignments)  
376.959 Million cell updates/sec

Title: US-09-912-628-7

Perfect score: 1145

Sequence: 1 HELRSWAARRTGAHRHGS.....SIPKNSPVTGHWXPPSL 215

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 908470

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1143	99.8	215	21	AA080776
2	1143	99.8	215	22	AA014914
3	896	78.3	199	21	AA080783
4	896	78.3	199	22	AA014921
5	893	78.0	202	21	AA080784
6	893	78.0	202	22	AA014922
7	885.5	77.3	201	21	AA080777
8	885.5	77.3	201	22	AA014915
9	600	52.4	170	23	ABG66743
10	404.5	35.3	319	22	AAU02922

11	404.5	35.3	372	22	AAU02963	Angiotensin conver
12	404.5	35.3	400	22	AAU02962	Angiotensin conver
13	403.5	35.2	204	22	AAU02974	Angiotensin conver
14	403.5	35.2	376	22	AA059176	Thrombin inhibitor
15	403.5	35.2	376	22	AA011125	Human thrombin inh
16	403.5	35.2	389	22	AAU02973	Angiotensin conver
17	403.5	35.2	459	21	AA056862	Angiotensin conver
18	399	34.8	156	22	AAU02975	Human prostate can
19	392.5	34.3	374	17	AA099253	Angiotensin conver
20	392.5	34.3	374	21	AA055840	Cytoplasmic antipr
21	392.5	34.3	374	21	AA055840	Novel cytoplasmic
22	320	27.9	376	17	AA099254	Cytoplasmic antipr
23	320	27.9	376	21	AA055841	Human cytoplasmic
24	278	24.3	390	18	AA015241	Psoriasis type
25	272	23.8	390	18	AA015242	Human SCCA2 protei
26	272	23.8	390	20	AA025927	Human SCCA2 protei
27	272	23.8	390	22	AA025927	Human SCCA2 protei
28	266	23.2	390	13	AA025276	Human squamous cel
29	266	23.2	390	20	AA025276	Human squamous cel
30	266	23.2	390	20	AA025276	Human squamous cel
31	266	23.2	390	20	AA025276	Human squamous cel
32	266	23.2	390	20	AA025276	Human squamous cel
33	266	23.2	390	20	AA025276	Human squamous cel
34	246	21.5	382	12	AA012144	Plasminogen activa
35	240	21.0	379	17	AA094367	Human elastase inh
36	239	20.9	437	21	AA043755	Human cancer assoc
37	235	20.5	379	16	AA064159	Human cancer assoc
38	232	20.3	392	12	AA012143	Human cancer assoc
39	229	20.0	41	21	AA080782	Angiotensin conver
40	229	20.0	41	21	AA080782	Angiotensin conver
41	225.5	19.7	438	12	AA014920	Plasminogen activa
42	225.5	19.7	438	12	AA014920	Plasminogen activa
43	224	19.6	224	22	AA072660	Human serine prote
44	224	19.6	316	20	AA028644	Human headpin (for
45	224	19.6	391	20	AA041029	Human lung tumor a
	224	19.6	391	20	AA025926	Human lung tumor a

## ALIGNMENTS

RESULT 1  
AA080776  
ID AAB08776 standard; Protein; 215 AA.  
XX AAB08776;  
XX AC  
XX AAB08776;  
DT 02-JAN-2001 (first entry)  
XX  
XX Amino acid sequence of a human serpin polypeptide.  
DE  
XX Human; serpin; serine protease inhibitor; autoimmune disease; neoplasia;  
KW rheumatoid arthritis; hyperproliferative disorder; cerebral ischemia;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW anglogenesis; nervous system disorder; Alzheimer's disease; infection;  
KW ocular disorder; corneal infection; epithelial cell proliferation;  
KW skin aging; sunburn; wound healing; organ transplantation; chemotaxis;  
KW food additive; food preservative.  
XX  
XX Homo sapiens.

OS  
XX Homo sapiens.  
FH Key Location/Qualifiers  
FT Misc-difference 211 /note= "any L-amino acid which is encoded by TGN"

XX  
XX WO200052160-A1.  
XX  
XX 08-SEP-2000.  
XX  
XX 29-FEB-2000; 2000WO-US05082.  
XX  
XX 01-MAR-1999; 99US-0122276.  
XX 12-MAR-1999; 99US-0124094.  
XX 18-AUG-1999; 99US-0149452.

XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Ruben SM, Ni J;  
XX PN WPI; 2000-579284/54.  
XX PD N-PSDB; AAA64690.  
XX PF Novel human serine protease inhibitor superfamily of proteins useful  
XX PT for diagnosing, treating, preventing and/or prognosing disorders  
XX PT related to the Serpin polypeptides  
XX PS Claim 12; Page 204; 215pp; English.  
XX PT The present sequence represents a human serpin polypeptide. Serpin is  
XX CC a member of the serine protease inhibitor superfamily of polypeptides.  
XX CC The serpin polynucleotides and polypeptides are used to prevent, treat  
XX CC or ameliorate a medical conditions in e.g. humans, mice, rabbits, goats,  
XX CC horses, cats, dogs, chickens or sheep. Disorders which are diagnosed or  
XX CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
XX CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
XX CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
XX CC e.g. cerebral ischemia, angiogenesis, nervous system disorders,  
XX CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
XX CC fungi and ocular disorders e.g. corneal infection. The polypeptides can  
XX CC also be used to aid wound healing and epithelial cell proliferation,  
XX CC to prevent skin aging due to sunburn, to maintain organs before  
XX CC transplantation, for supporting cell culture of primary tissues, to  
XX CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
XX CC as a food additive or preservative to increase or decrease storage  
XX CC capabilities.  
XX SQ Sequence 215 AA;  
Query Match 99.8%; Score 1143; DB 21; Length 215;  
Best Local Similarity 100.0%; Pred. No. 1.3e-118;  
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HELSWAARRTGAHRHGCISIRSKWHICIKPEKARGKOLKLIFFPMSISSALAMVFMG 60  
DB 1 HELSWAARRTGAHRHGCISIRSKWHICIKPEKARGKOLKLIFFPMSISSALAMVFMG 60  
QY 61 AKGNTAAQMSQALCFSGKIGGEDGDIHRGFQSLVAINRTDTEYVLRRTANGLFGEKSYDFL 120  
DB 61 AKGNTAAQMSQALCFSGKIGGEDGDIHRGFQSLVAINRTDTEYVLRRTANGLFGEKSYDFL 120  
QY 121 TGFTDSCGKIFYQATIKQLDFVNDTEKSTTRVNSWVADTKAWKIIQTSLSHLEEPGIASS 180  
DB 121 TGFTDSCGKIFYQATIKQLDFVNDTEKSTTRVNSWVADTKAWKIIQTSLSHLEEPGIASS 180  
QY 181 SCYKACLSQPLLVSIPKCNPSVTPHGMWXPSSL 215  
DB 181 SCYKACLSQPLLVSIPKCNPSVTPHGMWXPSSL 215  
RESULT 2  
AAO14914  
ID AAO14914 standard; Protein; 215 AA.  
XX AC AAO14914;  
XX AC AAO14914;  
XX DT 19-JUL-2002 (first entry)  
XX DE Human serine protease inhibitor (serpin) HMCIS41.  
XX KW Human; serine protease inhibitors; serpin; blood coagulation;  
XX KW proteinase-activated physiological process; gene therapy;  
XX KW antitense therapy; cancer; inflammation; neurological disease;  
XX KW immune disease; HMCIS41.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers

FT Misc-difference 211  
FT /label= Unknown  
FT /note= "Encoded by TGN"  
XX PN WO200155390-A1.  
XX PD 02-AUG-2001.  
XX PF 26-JAN-2001; 2001WO-US02484.  
XX PR 28-JAN-2000; 2000US-0178769.  
XX PR 29-FEB-2000; 2000WO-US05082.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PT Ni J, Ruben SM, Shi Y;  
XX PD WPI; 2001-457730/49.  
XX DR N-PSDB; AAL42746.  
XX PT Isolated serine protease inhibitor polypeptide for treating, preventing  
XX PT and/or prognosing cancers, inflammation, immune disorders and  
XX PT neurological diseases and also for testing and detection e.g. diagnosis  
XX PT -  
XX Claim 12; Page 276-277; 288pp; English.  
XX PT The invention comprises the amino acid and coding sequences of a group of  
XX CC novel human serine protease inhibitors (serpins). Serpins are a  
XX CC superfamily of single chain proteins that contain a conserved structure  
XX CC of approximately 370-420 amino acids and generally range between 50 and  
XX CC 100 kDa in molecular mass. The majority of serpins function as protease  
XX CC inhibitors and so are involved in the regulation of several proteinase-  
XX CC activated physiological processes (e.g. blood coagulation). The Serpin  
XX CC DNA and amino acid sequences of the invention are useful in the treatment  
XX CC (e.g. gene therapy, antitense therapy) and diagnosis of conditions such  
XX CC as cancer, inflammation, neurological diseases and immune diseases. The  
XX CC present amino acid sequence represents a human serpin protein of the  
XX CC invention.  
XX SQ Sequence 215 AA;  
Query Match 99.8%; Score 1143; DB 22; Length 215;  
Best Local Similarity 100.0%; Pred. No. 1.3e-118;  
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HELSWAARRTGAHRHGCISIRSKWHICIKPEKARGKOLKLIFFPMSISSALAMVFMG 60  
DB 1 HELSWAARRTGAHRHGCISIRSKWHICIKPEKARGKOLKLIFFPMSISSALAMVFMG 60  
QY 61 AKGNTAAQMSQALCFSGKIGGEDGDIHRGFQSLVAINRTDTEYVLRRTANGLFGEKSYDFL 120  
DB 61 AKGNTAAQMSQALCFSGKIGGEDGDIHRGFQSLVAINRTDTEYVLRRTANGLFGEKSYDFL 120  
QY 121 TGFTDSCGKIFYQATIKQLDFVNDTEKSTTRVNSWVADTKAWKIIQTSLSHLEEPGIASS 180  
DB 121 TGFTDSCGKIFYQATIKQLDFVNDTEKSTTRVNSWVADTKAWKIIQTSLSHLEEPGIASS 180  
QY 181 SCYKACLSQPLLVSIPKCNPSVTPHGMWXPSSL 215  
DB 181 SCYKACLSQPLLVSIPKCNPSVTPHGMWXPSSL 215  
RESULT 3  
AAB08783  
ID AAB08783 standard; Protein; 199 AA.  
XX AC AAB08783;  
XX XX 02-JAN-2001 (first entry)  
XX DE Amino acid sequence of a human serpin polypeptide.  
XX XX

KW Human; serpin; serine protease inhibitor; autoimmune disease; neoplasm;  
 KW rheumatoid arthritis; hyperproliferative disorder; cerebral ischemia;  
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
 KW ocular disorder; corneal infection; epithelial cell proliferation;  
 KW skin aging; sunburn; wound healing; organ transplantation; chemotaxis;  
 KW food additive; food preservative.  
 XX  
 OS Homo sapiens.

XX  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 195  
 FT /note= "any L-amino acid which is encoded by TGN"

XX  
 XX  
 PN WO200052160-A1.  
 XX  
 PD 08-SEP-2000.  
 XX

XX  
 XX 29-FEB-2000; 2000WO-US05082.  
 PR 01-MAR-1999; 99US-0122276.  
 PR 12-MAR-1999; 99US-0124094.  
 PR 18-AUG-1999; 99US-0149452.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ruben SM, Ni J;  
 XX  
 DR WPI; 2000-579284/54.  
 DR N-PSDB; AAA74934.

XX  
 XX Novel human serine protease inhibitor superfamily of proteins useful  
 PT for diagnosing, treating, preventing and/or prognosing disorders  
 PT related to the Serpin polypeptides  
 XX  
 PS Claim 12; Page 13; 215pp; English.  
 XX  
 CC The present sequence represents a human serpin polypeptide. Serpin is  
 CC a member of the serine protease inhibitor superfamily of polypeptides.  
 CC The serpin polynucleotides and polypeptides are used to prevent, treat  
 CC or ameliorate a medical conditions in e.g. humans, mice, rabbits, goats,  
 CC horses, cats, dogs, chickens or sheep. Disorders which are diagnosed or  
 CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischemia, angiogenesis, nervous system disorders,  
 CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
 CC fungi and ocular disorders e.g. corneal infection. The polypeptides can  
 CC also be used to aid wound healing and epithelial cell proliferation,  
 CC to prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities.  
 XX

SQ Sequence 199 AA;  
 Query Match 78.3%; Score 896; DB 21; Length 199;  
 Best Local Similarity 98.8%; Pred. No. 3.5e-91;  
 Matches 171; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 43 LIFPMSISSALAMVFMGAKGNTAAQMSQALCFKSGIGEDGDIHRGFQSLLVAINRTDTE 102  
 DB 27 LFSPMSISSALAMVFMGAKGNTAAQMSQALCFKSGIGEDGDIHRGFQSLLVAINRTDTE 86

QY 103 YVLTANGFLGKESYDFLTGFTDSCGKFGYQATIKQLDFVNDTEKSTTRVNSWADTKTAW 162  
 DB 87 YVLTANGFLGKESYDFLTGFTDSCGKFGYQATIKQLDFVNDTEKSTTRVNSWADTKTAW 146

QY 163 KIIOTSLHLEEPGIASSSCYCKACLSQLLVHSIPKCNISPTVPHGMWXPSSL 215  
 DB 147 KIIOTSLHLEEPGIASSSCYCKACLSQLLVHSIPKCNISPTVPHGMWXPSSL 199

RESULT 4  
 AAO14921

ID AAO14921 standard; Protein; 199 AA.

XX  
 AC AAO14921;

XX 19-JUL-2002 (first entry)

XX Human serine protease inhibitor (serpin) 1.

XX Human; serine protease inhibitors; serpin; blood coagulation;  
 KW proteinase-activated physiological process; gene therapy;  
 KW antiseptic therapy; cancer; inflammation; neurological disease;  
 KW immune disease.  
 XX  
 OS Homo sapiens.

XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 195  
 FT /label= Unknown  
 FT /note= "Encoded by TGN"

XX WO200155390-A1.

XX  
 PD 02-AUG-2001.

XX 26-JAN-2001; 2001WO-US02484.

XX 28-JAN-2000; 2000US-0178769.

XX 29-FEB-2000; 2000WO-US05082.

XX (HUMA-) HUMAN GENOME SCI INC.

XX  
 PI Ni J, Ruben SM, Shi Y;  
 XX  
 DR WPI; 2001-457730/49.  
 DR N-PSDB; AAL42747.

XX Isolated serine protease inhibitor polypeptide for treating, preventing  
 PT and/or prognosing cancers, inflammation, immune disorders and  
 PT neurological diseases and also for testing and detection e.g. diagnosis  
 PT  
 XX

PS Claim 12; Page 15; 288pp; English.

XX The invention comprises the amino acid and coding sequences of a group of  
 CC novel human serine protease inhibitors (serpins). Serpins are a  
 CC superfamily of single chain proteins that contain a conserved structure  
 CC of approximately 370-420 amino acids and generally range between 50 and  
 CC 100 kDa in molecular mass. The majority of serpins function as protease  
 CC inhibitors and so are involved in the regulation of several proteinase-  
 CC activated physiological processes (e.g. blood coagulation). The Serpin  
 CC DNA and amino acid sequences of the invention are useful in the treatment  
 CC (e.g. gene therapy, antiseptic therapy) and diagnosis of conditions such  
 CC as cancer, inflammation, neurological diseases and immune diseases. The  
 CC present amino acid sequence represents a human serpin protein of the  
 CC invention.  
 XX

SQ Sequence 199 AA;

Query Match 78.3%; Score 896; DB 22; Length 199;  
 Best Local Similarity 98.8%; Pred. No. 3.5e-91;  
 Matches 171; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 43 LIFPMSISSALAMVFMGAKGNTAAQMSQALCFKSGIGEDGDIHRGFQSLLVAINRTDTE 102

DB 27 LFSPMSISSALAMVFMGAKGNTAAQMSQALCFKSGIGEDGDIHRGFQSLLVAINRTDTE 86

QY 103 YVLTANGFLGKESYDFLTGFTDSCGKFGYQATIKQLDFVNDTEKSTTRVNSWADTKTAW 162

DB 87 YVLTANGFLGKESYDFLTGFTDSCGKFGYQATIKQLDFVNDTEKSTTRVNSWADTKTAW 146

QY 163 KIIOTSLHLEPGLIASSCYCKACLSQPLLVHSPKNSPVTPTGHWXPPSL 215  
DB 147 KIIOTSLHLEPGLIASSCYCKACLSQPLLVHSPKNSPVTPTGHWXPPSL 199

RESULT 5  
AAB08784  
ID AAB08784 standard; Protein; 202 AA.  
XX  
AC AAB08784;  
XX  
DT 02-JAN-2001 (first entry)  
XX

DE Amino acid sequence of a human serpin polypeptide.

XX Human; serpin; serine protease inhibitor; autoimmune disease; neoplasm;  
KW rheumatoid arthritis; hyperproliferative disorder; cerebral ischemia;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
KW ocular disorder; corneal infection; epithelial cell proliferation;  
KW skin aging; sunburn; wound healing; organ transplantation; chemotaxis;  
KW food additive; food preservative.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH Misc-difference 31 /note= "any L-amino acid which is encoded by NNN"  
FT Misc-difference 32 /note= "any L-amino acid which is encoded by NNN"  
FT Misc-difference 33 /note= "any L-amino acid which is encoded by NNN"  
FT Misc-difference 198 /note= "any L-amino acid which is encoded by NNN"  
FT Misc-difference 198 /note= "any L-amino acid which is encoded by TGN"

XX WO200052160-A1.

XX 08-SEP-2000.

XX 29-FEB-2000; 2000WO-US05082.

XX 01-MAR-1999; 99US-0122276.

XX 12-MAR-1999; 99US-0124094.

XX 18-AUG-1999; 99US-0149452.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Ni J;

XX WPI; 2000-579284/54.

XX N-PSDB; AAA74935.

XX Novel human serine protease inhibitor superfamily of proteins useful

XX for diagnosing, treating, preventing and/or prognosing disorders

XX related to the Serpin polypeptides

XX Claim 12; Page 13-14; 215pp; English.

XX The present sequence represents a human serpin polypeptide. Serpin is  
XX a member of the serine protease inhibitor superfamily of polypeptides.  
XX The serpin polynucleotides and polypeptides are used to prevent, treat  
XX or ameliorate a medical condition in e.g. humans, mice, rabbits, goats,  
XX horses, cats, dogs, chickens or sheep. Disorders which are diagnosed or  
XX treated include autoimmune diseases e.g. rheumatoid arthritis,  
XX hyperproliferative disorders e.g. neoplasms of the breast or liver,  
XX cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
XX e.g. cerebral ischemia, angiodysplasia, nervous system disorders,  
XX e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
XX fungi and ocular disorders e.g. corneal infection. The polypeptides can  
XX also be used to aid wound healing and epithelial cell proliferation,  
XX to prevent skin aging due to sunburn, to maintain organs before  
XX transplantation, for supporting cell culture of primary tissues, to  
XX regenerate tissues and in chemotaxis. The polypeptides can also be used

CC as a food additive or preservative to increase or decrease storage  
XX capabilities.  
SQ Sequence 202 AA;

Query Match 78.0%; Score 893; DB 21; Length 202;  
Best Local Similarity 93.0%; Pred. No. 7.7e-91;  
Matches 173; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 34 KARGKQLKQLIFF-PMSSISALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGF 89  
DB 17 KLGNNNSNLFSSXXPMSSISALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGF 76  
QY 90 QSLLVAINRTDTEYVLRNTANGLFGEKSYDPLTGTDTSCGKFYQATIKQLDFVNDTEKSTT 149  
DB 77 QSLLVAINRTDTEYVLRNTANGLFGEKSYDPLTGTDTSCGKFYQATIKQLDFVNDTEKSTT 136

QY 150 RVNSWADKTKANKIQTSLHLEPGLIASSCYCKACLSQPLLVHSPKNSPVTPTGHWX 209  
DB 137 RVNSWADKTKANKIQTSLHLEPGLIASSCYCKACLSQPLLVHSPKNSPVTPTGHWX 196

QY 210 WXPPSL 215  
DB 197 WXPPSL 202

RESULT 6  
AAO14922  
ID AAO14922 standard; Protein; 202 AA.

AC AAO14922;

XX 19-JUL-2002 (first entry)

XX Human serine protease inhibitor (serpin) 2.

XX Human; serine protease inhibitors; serpin; blood coagulation;  
KW proteinase-activated physiological process; gene therapy;  
KW antitense therapy; cancer; inflammation; neurological disease;  
KW immune disease.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH Misc-difference 31 /label= Unknown  
FT Misc-difference 31 /note= "Encoded by NNN"  
FT Misc-difference 32 /label= Unknown  
FT Misc-difference 32 /note= "Encoded by NNN"  
FT Misc-difference 33 /label= Unknown  
FT Misc-difference 198 /note= "Encoded by NNN"  
FT Misc-difference 198 /label= Unknown  
FT Misc-difference 198 /note= "Encoded by TGN"

XX WO200155390-A1.

XX 02-AUG-2001.

XX 26-JAN-2001; 2001WO-US02484.

XX 28-JAN-2000; 2000US-0178769.

XX 29-FEB-2000; 2000WO-US05082.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Ruben SM, Shi Y;

XX WPI; 2001-457730/49.

XX N-PSDB; AAL42748.

PT Isolated serine protease inhibitor polypeptide for treating, preventing  
 PT and/or prognosing cancers, inflammation, immune disorders and  
 PT neurological diseases and also for testing and detection e.g. diagnosis  
 PT  
 XX  
 PS Claim 12; Page 15; 288pp; English.  
 PS  
 CC The invention comprises the amino acid and coding sequences of a group of  
 CC novel human serine protease inhibitors (serpins). Serpins are a  
 CC superfamily of single chain proteins that contain a conserved structure  
 CC of approximately 370-420 amino acids and generally range between 50 and  
 CC 100 kDa in molecular mass. The majority of serpins function as protease-  
 CC inhibitors and so are involved in the regulation of several proteinase-  
 CC activated physiological processes (e.g. blood coagulation). The Serpin  
 CC DNA and amino acid sequences of the invention are useful in the treatment  
 CC (e.g. gene therapy, antisense therapy) and diagnosis of conditions such  
 CC as cancer, inflammation, neurological diseases and immune diseases. The  
 CC present amino acid sequence represents a human serpin protein of the  
 CC invention.

SQ Sequence 202 AA;

Query Match 78.0%; Score 893; DB 22; Length 202;  
 Best Local Similarity 93.0%; Pred. No. 7.7e-91;  
 Matches 173; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 34 KARGKQLKLIFF----PMSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGF 89  
 DB 17 KKLGENNSNNLFFSXXPMSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGF 76  
 QY 90 QSLVAINRTDTEYVLTANGLFGEKSYDFTLTGFTDSCGKFYQATIKQLDFVNDTEKSTT 149  
 DB 77 QSLVAINRTDTEYVLTANGLFGEKSYDFTLTGFTDSCGKFYQATIKQLDFVNDTEKSTT 136  
 QY 150 RVNSWVADTKAWKIQTSLSHLEPGIASSSCYKACLSQPLLVHSPKCNPTVPHGM 209  
 DB 137 RVNSWVADTKAWKIQTSLSHLEPGIASSSCYKACLSQPLLVHSPKCNPTVPHGM 196  
 QY 210 WXPSSL 215  
 DB 197 WXPSSL 202

RESULT 7  
 AAB08777  
 ID AAB08777 standard; Protein; 201 AA.

Y  
 AAB08777;

DT 02-JAN-2001 (first entry)

DE Amino acid sequence of a human serpin polypeptide.

XX Human; serpin; serine protease inhibitor; autoimmune disease; neoplasm;  
 KW rheumatoid arthritis; hyperproliferative disorder; cerebral ischemia;  
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
 KW ocular disorder; corneal infection; epithelial cell proliferation;  
 KW skin aging; sunburn; wound healing; organ transplantation; chemotaxis;  
 KW food additive; food preservative.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 32 /note= "any L-amino acid"

FT Misc-difference 33 /note= "any L-amino acid"

FT Misc-difference 197 /note= "any L-amino acid"

FT Misc-difference 197 /note= "any L-amino acid"

PN WO200052160-A1.

XX

PD 08-SEP-2000.  
 XX 29-FEB-2000; 2000WO-US05082.  
 PF 01-MAR-1999; 99US-0122276.  
 PR 12-MAR-1999; 99US-0124094.  
 PR 18-AUG-1999; 99US-0149452.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Ruben SM, Ni J;  
 XX WPI; 2000-579284/54.

DR Novel human serine protease inhibitor superfamily of proteins useful  
 XX for diagnosing, treating, preventing and/or prognosing disorders  
 XX related to the Serpin polypeptides -  
 XX Claim 12; Page 12; 215pp; English.

XX The present sequence represents a human serpin polypeptide. Serpin is  
 CC a member of the serine protease inhibitor superfamily of polypeptides.  
 CC The serpin polynucleotides and polypeptides are used to prevent, treat  
 CC or ameliorate a medical conditions in e.g. humans, mice, rabbits, goats,  
 CC horses, cats, dogs, chickens or sheep. Disorders which are diagnosed or  
 CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischemia, angiogenesis, nervous system disorders,  
 CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
 CC fungi and ocular disorders e.g. corneal infection. The polypeptides can  
 CC also be used to aid wound healing and epithelial cell proliferation,  
 CC to prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities.

XX Query Match 77.3%; Score 885.5; DB 21; Length 201;  
 Best Local Similarity 93.0%; Pred. No. 5.2e-90;  
 Matches 172; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY 34 KARGKQLKLIFFP--MSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFQ 90  
 DB 17 KKLGENNSNNLFFSXXMSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFQ 76  
 QY 91 SLLVAINRTDTEYVLTANGLFGEKSYDFTLTGFTDSCGKFYQATIKQLDFVNDTEKSTTR 150  
 DB 77 SLLVAINRTDTEYVLTANGLFGEKSYDFTLTGFTDSCGKFYQATIKQLDFVNDTEKSTTR 136  
 QY 151 VNSWVADTKAWKIQTSLSHLEPGIASSSCYKACLSQPLLVHSPKCNPTVPHGMW 210  
 DB 137 VNSWVADTKAWKIQTSLSHLEPGIASSSCYKACLSQPLLVHSPKCNPTVPHGMW 196  
 QY 211 XPPSL 215  
 DB 197 XPPSL 201

SQ Sequence 201 AA;

Query Match 77.3%; Score 885.5; DB 21; Length 201;  
 Best Local Similarity 93.0%; Pred. No. 5.2e-90;  
 Matches 172; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY 34 KARGKQLKLIFFP--MSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFQ 90  
 DB 17 KKLGENNSNNLFFSXXMSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFQ 76  
 QY 91 SLLVAINRTDTEYVLTANGLFGEKSYDFTLTGFTDSCGKFYQATIKQLDFVNDTEKSTTR 150  
 DB 77 SLLVAINRTDTEYVLTANGLFGEKSYDFTLTGFTDSCGKFYQATIKQLDFVNDTEKSTTR 136  
 QY 151 VNSWVADTKAWKIQTSLSHLEPGIASSSCYKACLSQPLLVHSPKCNPTVPHGMW 210  
 DB 137 VNSWVADTKAWKIQTSLSHLEPGIASSSCYKACLSQPLLVHSPKCNPTVPHGMW 196  
 QY 211 XPPSL 215  
 DB 197 XPPSL 201

RESULT 8  
 AAO14915  
 ID AAO14915 standard; Protein; 201 AA.  
 XX  
 AC AAO14915;  
 XX  
 DT 19-JUL-2002 (first entry)  
 XX  
 DE Human serine protease inhibitor (serpin) 3.  
 XX Human; serine protease inhibitors; serpin; blood coagulation;  
 KW proteinase-activated physiological process; gene therapy;

antiseize therapy; cancer; inflammation; neurological disease;  
immune disease.

Homo sapiens.

Key Location/Qualifiers  
Misc-difference 32  
/label= Unknown  
Misc-difference 33  
/label= Unknown  
Misc-difference 197  
/label= Unknown

WO200155390-A1.

02-AUG-2001.

26-JAN-2001; 2001WO-US02484.

28-JAN-2000; 2000US-0178769.

29-FEB-2000; 2000WO-US05082.

(HUMA-) HUMAN GENOME SCI INC.

Ni J, Ruben SM, Shi Y;

WPI; 2001-457730/49.

Isolated serine protease inhibitor polypeptide for treating, preventing  
and/or prognosing cancers, inflammation, immune disorders and  
neurological diseases and also for testing and detection e.g. diagnosis

Claim 12; Page 14; 288pp; English.

The invention comprises the amino acid and coding sequences of a group of  
novel human serine protease inhibitors (serpins). Serpins are a  
superfamily of single chain proteins that contain a conserved structure  
of approximately 370-420 amino acids and generally range between 50 and  
100 kDa in molecular mass. The majority of serpins function as protease  
inhibitors and so are involved in the regulation of several proteinase-  
activated physiological processes (e.g. blood coagulation). The Serpin  
DNA and amino acid sequences of the invention are useful in the treatment  
(e.g. gene therapy, antisense therapy) and diagnosis of conditions such  
as cancer, inflammation, neurological diseases and immune diseases. The  
present amino acid sequence represents a human serpin protein of the  
invention.

Sequence 201 AA;

Query Match 77.3%; Score 885.5; DB 22; Length 201;  
Best Local Similarity 93.0%; Pred. No. 5.2e-90;  
Matches 172; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY 34 KARGKQLKLIFFP---MSISSALAMVFMGAKGNTAAQMSQALCFKIGGEDGDIHRGQ 90

Db 17 KLGNNNNLFFSPXMSISSALAMVFMGAKGNTAAQMSQALCFKIGGEDGDIHRGQ 76

QY 91 SLIVAINRTDTEVLTANGLFGEKSYDFLTGTDSCGKGYQNTIKQLDFVNDTEKSTTR 150

Db 77 SLIVAINRTDTEVLTANGLFGEKSYDFLTGTDSCGKGYQNTIKQLDFVNDTEKSTTR 136

QY 151 VNSWADTKAWKIQTSLSHLEPGFASSSCYCKACLSPILLVHSIPKCNPSVTPHGMW 210

Db 137 VNSWADTKAWKIQTSLSHLEPGFASSSCYCKACLSPILLVHSIPKCNPSVTPHGMW 196

QY 211 XPPSL 215

Db 197 XPPSL 201

RESULT 9  
ABG66743

ABG66743 standard; Protein; 170 AA.

ABG66743;

30-AUG-2002 (first entry)

Human novel polypeptide #78.

Human; inflammatory condition; shock; sepsis; immune response;  
cancer; wound healing; central nervous system disease; haematopoiesis;  
peripheral nervous system disease; ankyrotrophic lateral sclerosis; tendon;  
myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;  
cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;  
bone degenerative disorder; periodontal disease; reperfusion injury;  
lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;  
allergic condition; thrombolysis; thrombosis; coagulation disorder;  
fungal infection.

Homo sapiens.

WO200244340-A2.

06-JUN-2002.

30-NOV-2001; 2001WO-US47004.

30-NOV-2000; 2000US-0028952.

(HYSE-) HYSEQ INC.

Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;  
Yamazaki V, Ujwal ML, Drmanac RT;

WPI; 2002-508509/54.

N-PSDB; ABK94967.

Novel nucleic acids and polypeptides for diagnosis, treatment of  
inflammatory, autoimmune, nervous system, myeloid or lymphoid cell  
disorders, cancer and promoting wound healing -

Claim 10; Page 654; 672pp; English.

The invention relates to human novel polynucleotides and associated  
polypeptides. The polynucleotides and polypeptides are useful for  
treating inflammatory conditions such as arthritis, nephritis, Crohn's  
disease, ischaemia-reperfusion injury, shock, sepsis, immune responses  
and cancer and for promoting wound healing. The sequences are used to  
induce the proliferation of neural cells and regeneration of nerve and  
brain tissue, and are useful for the treatment of central and peripheral  
nervous system diseases and neuropathies, such as Alzheimer's disease,  
Parkinson's disease, Huntington's disease and amyotrophic lateral  
sclerosis. The sequences are involved in chemotactic or chemokinetic  
activity, regulation of haematopoiesis, treatment of myeloid or lymphoid  
cell disorders and platelet disorders such as thrombocytopenia,  
regeneration of bone, cartilage, tendon, ligament and/or nerve tissue  
growth, tissue repair, healing of burns, incisions, ulcers, treatment of  
osteoporosis, osteoarthritis, bone degenerative disorders and periodontal  
disease. The sequences of the invention are also useful for gut  
protection or regeneration and treatment of lung or liver fibrosis,  
reperfusion injury in various tissues, immune deficiencies and disorders  
including severe combined immunodeficiency (SCID), bacterial or fungal  
infections, autoimmune disorders e.g. multiple sclerosis and myasthenia  
gravis, allergic conditions such as asthma, thrombolysis or thrombosis  
and coagulation disorders. Sequences ABG66666-ABG66758 represent human  
novel polypeptides of the invention.

Sequence 170 AA;

Query Match 52.4%; Score 600; DB 23; Length 170;

Best Local Similarity 83.9%; Pred. No. 2.4e-58;

Matches 120; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

QY 43 LIFFPMSISSALAMVFMGAKGNTAAQMSQALCFKIGGEDGDIHRGFSILVAINRTDTE 102



CC	abnormality such as deep vein thrombosis.
XX	
QQ	Sequence 319 AA;
Query Match	35.3%; Score 404.5; DB 22; Length 319;
Best Local Similarity	51.5%; Pred. No. 3.3e-36;
Matches 85; Conservative 29; Mismatches 36; Indels 15; Gaps	
QY	34 KARGKQLKLIFF-PMSISSALAMVFWGAKGNTAAQNSOALCFCKTGGEDGDIHRGFQSL 92
DB	17 KTLGKDNSKNVFFSPMSKSCALAMVYWGAKGNTAAQNAQILSFNKGSG-GGDIHQGFQSL 75
QY	93 LVAINRTDEYVLRTRANGLFGEKGYDFLTGFTDSCGKGYQATIKQLDFVNDTEKSTRW 152
DB	76 LTEVNTKTGYLLRVANFLFGEKSCDFLSSFRSCQFYQAEHEEDLFISAVEKSRKHIN 135
QY	153 SWVADKTKAWKIQTSLSHLEEPGIASSCYKACLSQPLLVHSI 197
DB	136 TWVAEKTEG-----KIAELLSPGSVDP-----LTRLVLVNAV 167
RESULT 11	
AAU02963	
ID	AAU02963 standard; Protein; 372 AA.
AC	AAU02963;
XX	
DT	12-SEP-2001 (first entry)
XX	
DE	Angiotensin converting enzyme (ACEV) splice variant protein #63.
XX	
KW	Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW	granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW	platelet-derived endothelial cell growth factor; cardiovascular disease
KW	cellular tumour antigen p53; cyclin-dependent kinase inhibitor 1C;
KW	vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW	myocardial infarction; coronary arterial thrombosis; renal disease;
KW	diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW	multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW	noncardiogenic pulmonary granulomatous disease; endothelial abnormality;
KW	vascular disorder; asbestosis.
OS	Homo sapiens.
XX	
PN	WO200136632-A2.
XX	
PD	25-MAY-2001.
XX	
PF	17-NOV-2000; 2000WO-IL00766.
XX	
PR	17-NOV-1999; 99IL-0132978.
XX	
PR	10-DEC-1999; 99IL-0133455.
XX	
PA	(COMP-) COMPUGEN LTD.
XX	
PI	Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX	
DR	WPI; 2001-336004/35.
XX	
DR	N-PSDB; AAS06063.
XX	
PFT	Novel alternative splicing variants e.g. variant of angiotensin
PFT	converting enzyme (ACEV), useful in identifying candidate compounds
PFT	capable of binding to the variant and to detect anti-variant antibodies
PFT	
PPS	Claim 4; Fig 63; 519pp; English.
XX	
CC	The sequence represents an angiotensin converting enzyme splice variant
CC	(ACEV) polypeptide. The polypeptides of the invention include variants
CC	granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC	platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC	inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal
CC	polypeptide receptor 2. The polypeptides and their associated nucleic

CC acids are useful for identification of variant sequences and detection of  
CC candidate compounds capable of binding the molecules. The sequences of  
CC the invention can be used in the treatment and diagnosis of various  
CC disorders including cardiovascular diseases such as arteriosclerosis,  
CC myocardial infarction and coronary arterial thrombosis, renal diseases  
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,  
CC immune disorders such as immune complex nephritis, multiple sclerosis,  
CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such  
CC as asbestosis and vascular pathologies involving an endothelial  
CC abnormality such as deep vein thrombosis.

XX Sequence 372 AA;  
SQ  
Query Match 35.3%; Score 404.5; DB 22; Length 372;  
Best Local Similarity 51.5%; Pred. No. 4.2e-36;  
Matches 85; Conservative 29; Mismatches 36; Indels 15; Gaps 4;  
QY 34 KARGKQLQLIFF-PMSISSALAMVFMGAKNTAAQMSQALCFKSGEDGDIHRGFSQL 92  
DB 17 KTLGKDNKSNVFFSPMSALAMVFMGAKNTAAQMSQALCFKSGEDGDIHRGFSQL 75  
QY 93 LVAINRTDTEYVLTANGLFGEKSYDFLTGFTDSCGKFYQATIKQLDFVNDTEKSTRVN 152  
DB 76 LTVNKGTQYLLRVANRLFGKSCDFLSFRDSCQKFYQAEEMELDFISAVEKSRKHIN 135  
QY 153 SWADKTKAWKIIQTSLSHLEPGIASSSCYCKACLSQPLLVHSI 197  
DB 136 TWAKEGTEG-----KIAELLSPGSVDP-----LTRLVLVNAV 167  
RESULT 12  
AAU02962  
ID AAU02962 standard; Protein; 400 AA.  
XX  
AC AAU02962;  
XX  
DT 12-SEP-2001 (first entry)  
XX  
DE Angiotensin converting enzyme (ACEV) splice variant protein #62.  
XX  
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;  
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;  
KW platelet-derived endothelial cell growth factor; cardiovascular disease;  
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;  
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;  
KW myocardial infarction; coronary arterial thrombosis; renal disease;  
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;  
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;  
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;  
KW vascular disorder; asbestosis.

OS Homo sapiens.  
XX  
FN WO200136632-A2.  
XX  
PD 25-MAY-2001.  
XX  
PF 17-NOV-2000; 2000WO-IL00766.  
XX  
PR 17-NOV-1999; 99IL-0132978.  
PR 10-DEC-1999; 99IL-0133455.  
XX  
PA (COMP-) COMPUEN LTD.  
XX  
PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;  
XX  
DR WPI; 2001-336004/35.  
DR N-PSDB; AAS06062.  
XX  
PT Novel alternative splicing variants e.g. variant of angiotensin  
PT converting enzyme (ACEV), useful in identifying candidate compounds  
PT capable of binding to the variant and to detect anti-variant antibodies  
PT

XX Claim 4; Fig 62; 519pp; English.  
PS  
XX The sequence represents an angiotensin converting enzyme splice variant  
XX (ACEV) polypeptide. The polypeptides of the invention include variants of  
XX granulocyte colony stimulating factor receptor, glucagon, interleukin 6,  
XX platelet-derived endothelial cell growth factor, cyclin-dependent kinase  
XX inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal  
XX polypeptide receptor 2. The polypeptides and their associated nucleic  
XX acids are useful for identification of variant sequences and detection of  
XX candidate compounds capable of binding the molecules. The sequences of  
XX the invention can be used in the treatment and diagnosis of various  
XX disorders including cardiovascular diseases such as arteriosclerosis,  
XX myocardial infarction and coronary arterial thrombosis, renal diseases  
XX such as diabetic nephropathy, muscular diseases such as hypertrophy,  
XX immune disorders such as immune complex nephritis, multiple sclerosis,  
XX cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such  
XX as asbestosis and vascular pathologies involving an endothelial  
XX abnormality such as deep vein thrombosis.

SQ Sequence 400 AA;  
Query Match 35.3%; Score 404.5; DB 22; Length 400;  
Best Local Similarity 51.5%; Pred. No. 4.6e-36;  
Matches 85; Conservative 29; Mismatches 36; Indels 15; Gaps 4;  
QY 34 KARGKQLQLIFF-PMSISSALAMVFMGAKNTAAQMSQALCFKSGEDGDIHRGFSQL 92  
DB 17 KTLGKDNKSNVFFSPMSALAMVFMGAKNTAAQMSQALCFKSGEDGDIHRGFSQL 75  
QY 93 LVAINRTDTEYVLTANGLFGEKSYDFLTGFTDSCGKFYQATIKQLDFVNDTEKSTRVN 152  
DB 76 LTVNKGTQYLLRVANRLFGKSCDFLSFRDSCQKFYQAEEMELDFISAVEKSRKHIN 135  
QY 153 SWADKTKAWKIIQTSLSHLEPGIASSSCYCKACLSQPLLVHSI 197  
DB 136 TWAKEGTEG-----KIAELLSPGSVDP-----LTRLVLVNAV 167

RESULT 13  
AAU02974  
ID AAU02974 standard; Protein; 204 AA.  
XX  
AC AAU02974;  
XX  
DT 12-SEP-2001 (first entry)  
XX  
DE Angiotensin converting enzyme (ACEV) splice variant protein #74.  
XX  
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;  
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;  
KW platelet-derived endothelial cell growth factor; cardiovascular disease;  
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;  
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;  
KW myocardial infarction; coronary arterial thrombosis; renal disease;  
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;  
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;  
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;  
KW vascular disorder; asbestosis.  
XX  
OS Homo sapiens.  
XX  
FN WO200136632-A2.  
XX  
PD 25-MAY-2001.  
XX  
PF 17-NOV-2000; 2000WO-IL00766.  
XX  
PR 17-NOV-1999; 99IL-0132978.  
PR 10-DEC-1999; 99IL-0133455.  
XX  
PA (COMP-) COMPUEN LTD.  
XX

PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;  
 XX WPI; 2001-336004/35.  
 DR N-PSDB; AAS06074.  
 XX  
 PT Novel alternative splicing variants e.g. variant of angiotensin  
 PT converting enzyme (ACEV), useful in identifying candidate compounds  
 PT capable of binding to the variant and to detect anti-variant antibodies  
 XX  
 PS Claim 4; Fig 74; 519pp; English.  
 XX  
 CC The sequence represents an angiotensin converting enzyme splice variant  
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of  
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,  
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase  
 CC inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal  
 CC polypeptide receptor 2. The polypeptides and their associated nucleic  
 CC acids are useful for identification of variant sequences and detection of  
 CC candidate compounds capable of binding to the molecules. The sequences of  
 CC the invention can be used in the treatment and diagnosis of various  
 CC disorders including cardiovascular diseases such as arteriosclerosis,  
 CC myocardial infarction and coronary arterial thrombosis, renal diseases  
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,  
 CC immune disorders such as immune complex nephritis, multiple sclerosis,  
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such  
 CC as asbestosis and vascular pathologies involving an endothelial  
 CC abnormality such as deep vein thrombosis.

XX SQ Sequence 204 AA;  
 Query Match 35.2%; Score 403.5; DB 22; Length 204;  
 Best Local Similarity 51.5%; Pred. No. 2.2e-36;  
 Matches 85; Conservative 29; Mismatches 36; Indels 15; Gaps 4;  
 QY 34 KARGKQLQLIFF-PMSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFSQL 92  
 Db 17 KTLGKDNKNVFFSPMSMSCALAMVYMGAKGNTAAQMAQILSFNKS GG-GGDIHQGFQSL 75  
 QY 93 LVAINRTDEYVLTANGLFGEKSYDPLTGFTDSCGKPYQATIKQLDFVNDTEKSTTRVN 152  
 Db 76 LTVNKTQYLLRMANRLFGEKSCDFLSPRDSQKPYQAEEMELDFISAVEKSRKHIN 135  
 QY 153 SWADKTKAWKIQTSLSHLEBPFGIASSSCYCKACLSQPLLVHSI 197  
 Db 136 TWAEKTEG-----KIAELSPGSVDP-----LTRLVLVNAV 167

RESULT 14  
 AAB59176  
 ID AAB59176 standard; Protein; 376 AA.  
 AC AAB59176;  
 XX  
 DT 22-MAR-2001 (first entry)  
 XX  
 DE Thrombin inhibitor protein.  
 XX  
 KW Thrombin inhibitor; neurodegenerative disease; Alzheimer's;  
 KW Parkinson's; tumour; deep venous thrombosis; pulmonary embolism;  
 KW heart.  
 XX  
 OS Unidentified.  
 XX  
 PN US6156540-A.  
 XX  
 PD 05-DEC-2000.  
 XX  
 PF 22-DEC-1993; 93US-0171817.  
 XX  
 PR 22-DEC-1993; 93US-0171817.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Fuldner RA, Cao L, Adams MD;  
 XX WPI; 2001-060165/07.  
 XX  
 PT Novel polynucleotide encoding thrombin inhibitor polypeptide useful for  
 PT diagnosis and treatment of diseases related to thrombosis -  
 XX  
 PS Claim 1; Fig 1; 19pp; English.  
 XX  
 CC The present invention relates to a thrombin inhibitor. This is useful  
 CC for diagnosis and treatment of thrombosis, neurodegenerative diseases  
 CC such as Alzheimer's disease and Parkinson's disease, tumor metastasis,  
 CC to prevent proximal extension of deep venous thrombosis or the  
 CC recurrence of pulmonary embolisms and also to treat risk patients who  
 CC have congestive heart failure, acute myocardial infarction or  
 CC cardiomyopathy to prevent the development of deep vein thrombosis or  
 CC pulmonary embolism. The thrombin inhibitor is also useful for  
 CC screening chemical compounds that either up or down regulate its  
 CC activity.  
 XX SQ Sequence 376 AA;

Query Match 35.2%; Score 403.5; DB 22; Length 376;  
 Best Local Similarity 51.5%; Pred. No. 5.5e-36;  
 Matches 85; Conservative 29; Mismatches 36; Indels 15; Gaps 4;  
 QY 34 KARGKQLQLIFF-PMSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFSQL 92  
 Db 17 KTLGKDNKNVFFSPMSMSCALAMVYMGAKGNTAAQMAQILSFNKS GG-GGDIHQGFQSL 75  
 QY 93 LVAINRTDEYVLTANGLFGEKSYDPLTGFTDSCGKPYQATIKQLDFVNDTEKSTTRVN 152  
 Db 76 LTVNKTQYLLRMANRLFGEKSCDFLSPRDSQKPYQAEEMELDFISAVEKSRKHIN 135  
 QY 153 SWADKTKAWKIQTSLSHLEBPFGIASSSCYCKACLSQPLLVHSI 197  
 Db 136 TWAEKTEG-----KIAELSPGSVDP-----LTRLVLVNAV 167

RESULT 15  
 AAB11125  
 ID AAB11125 standard; Protein; 376 AA.  
 AC AAB11125;  
 XX  
 DT 16-FEB-2001 (first entry)  
 XX  
 DE Human thrombin inhibitor protein.  
 XX  
 KW Thrombin inhibitor; human; thrombolytic; cardiant; cytostatic; nootropic;  
 KW neuroprotective; antiparkinsonian; gene therapy; treatment; carcinoma;  
 KW cell hypercoagulation; deep venous thrombosis; pulmonary embolism;  
 KW cerebral embolism; thrombic disease; tumor metastases; clot accretion;  
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6133422-A.  
 XX  
 PD 17-OCT-2000.  
 XX  
 PF 30-NOV-1998; 98US-0200965.  
 XX  
 PR 22-DEC-1993; 93US-0171817.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Fuldner RA, Rosen CA, Cao L, Adams MD;  
 XX WPI; 2001-006150/01.  
 DR N-PSDB; AAC65737.  
 XX

Wed May 28 14:44:09 2003

PT New thrombin inhibiting proteins for treating or preventing  
PT thrombosis-related diseases, e.g. embolism, tumor metastases,  
PT carcinoma, neurodegenerative diseases, e.g. Alzheimer's or Parkinson's  
PT disease -  
XX  
XX  
PS Claim 2; Column 17-20; 19pp; English.  
XX  
CC This invention describes a novel human thrombin inhibitor which has  
CC thrombolytic, cardiant, cytostatic, neurotropic, neuroprotective and  
CC antiparkinsonian activity and can be used for gene therapy. The thrombin  
CC inhibitor may be used therapeutically or diagnostically to treat or  
CC prevent diseases related to thrombosis characterized by hypercoagulation  
CC of cells. The protein may also be used to prevent extension of deep  
CC venous thrombosis of the recurrence of pulmonary embolisms and  
CC recurrence of cerebral or other systemic embolisms, to treat high risk  
CC patients, e.g. those who have congestive heart failure, acute myocardial  
CC infarction or cardiomyopathy to prevent the development of deep vein  
CC thrombosis or pulmonary embolism, and as long term therapy for occasional  
CC patient who has recurrent thrombosis or embolism while on the drug  
CC warfarin. The protein may also be used for treating thrombotic disease,  
CC tumor metastases, carcinoma of the brain, liver, lung, bone and  
CC neoplastic plasma cell carcinoma, and neurodegenerative diseases (e.g.  
CC Alzheimer's or Parkinson's disease). It can also be used against  
CC clot-bound thrombin, which produces clot accretion, for screening  
CC chemical compounds that either up or down regulate its activity, and for  
CC producing antibodies. The thrombin inhibitor can be used as  
CC pharmaceutical to mammals, such as humans, mice, rats, hamster, dogs,  
CC rabbits and cats. The cDNA sequences of the thrombin inhibitor may be  
CC used as diagnostic probe, for chromosome identification.

XX SQ Sequence 376 AA;

Query Match 35.2%; Score 403.5; DB 22; Length 376;  
Best Local Similarity 51.5%; Pred. No. 5.5e-36;  
Matches 85; Conservative 29; Mismatches 36; Indels 15; Gaps 4;  
QY 34 KARGKQLKQLIFF-PMSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFQSL 92  
Db 17 KTLGKDNSKNVFFSPMSMSALAMVFMGAKGNTAAQMAQLISFNKSGG-GGDIHCGFQSL 75  
QY 93 LVAINRTDTEYVLRNANGLFGEKSYDFLTGFTDSCGKFYQATIKOLDVFNDEKSTTRVN 152  
Db 76 LTVNKVTGTQYLLRWANRIFGEKSCDFLSSFRDSCQKFYQAEEMELDFISAVEKSRKHIN 135  
QY 153 SWVADKTKAWKLIQISLHLEPGIASSSCYCKACLSQELLVHSI 197  
Db 136 TWVAEXTG-----KIAELLSPGSVDP-----LTRLVLVNAV 167

Search completed: May 20, 2003, 10:24:01  
Time : 78 secs



Wed May 28 14:44:09 2003

```
Db 17 KTLGKNSKNVFPSPMSALAMVFMGAKNTAAQMAQILSPFNKSGG-GGDIHQGFQSL 75
Qy 93 LVAINRTDTEVLTANGLFGKSYDFTGTDSCGKGYQATIKQLDFVNDTEKSTTRVN 152
Db 76 LTVNKGTGYLLRANLFGKSCDFLSPFSCQKGYQAEWELDFISAVEKSRKHIN 135
Qy 153 SWADKTKAWKIIQTSLSHLEEPGIASSCCYCKACLSQPLLVHSI 197
Db 136 TWAEKTEG-----KIAELSPGSVDP-----LTRLVLVNAV 167

RESULT 2
US-08-464-148-2
; Sequence 2, Application US/08464148
; Patent No. 5710026
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,500
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 13952-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-385-500-2

Query Match 34.3%; Score 392.5; DB 1; Length 374;
Best Local Similarity 52.9%; Pred. No. 9e-37;
Matches 83; Conservative 23; Mismatches 34; Indels 17; Gaps 3;

Qy 41 KQIIFPMSISSALAMVFMGAKNTAAQMSQALCFSGKIGGEDGDIHRGFQSLLVAINRTD 100
Db 26 RNVEFPMSISSALAMVFMGAKNTAAQMSQALCLYK-----DGDIIHRGFQSLLVSEVNR 81
Qy 101 TEVLTANGLFGKSYDFTGTDSCGKGYQATIKQLDFVNDTEKSTTRVNSWADKTK 160
Db 82 TQYLLRTANLFGKTCDFLPDFKEYCQKGYQAELELSFAEDTEECRKHINDWVAEKTE 141
Qy 161 AWKIIQTSLSHLEEPGIASSCCYCKACLSQPLLVHSI 197
Db 142 G-----KISEVLDA GTVP-----LTKLVLVNAV 165

RESULT 4
US-08-846-784-2
; Sequence 2, Application US/08846784
; Patent No. 5747645
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/846,784  
FILING DATE: 30-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/385,500  
FILING DATE: 08-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 13952-21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 374 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-846-784-2

Query Match 34.3%; Score 392.5; DB 1; Length 374;  
Best Local Similarity 52.9%; Pred. No. 9e-37;  
Matches 83; Conservative 23; Mismatches 34; Indels 17; Gaps 3;  
QY 41 KOLIFFPMSISSALAMVFMGAKGNTAAQMSQALCFKSGIGGEDGDIHRGFQSLLVAINRTD 100  
DB 26 RNVFSPMSISSALAMVFMGAKGNTAAQMSQALCLYK---DGDHHRGFQSLLVSEVNRGTG 81  
QY 101 TEYVLTANGLFGEKSYDFTGTDSCGKEYQATIKOLDVNDTEKSTTRVNSWVADTK 160  
DB 82 TQYLLRTANRLFGKTCDFLPDFKEYCKFKYQAELEELSPFADTECKRHNDWVAEKTE 141  
QY 161 AKWIIQTSLSHLEPFGIASSCYCKACLSQPLLVHSI 197  
DB 142 G-----KISEVLDAAGTVDP-----LTKLVLVNAI 165

RESULT 5  
US-08-464-148-4  
Sequence 4, Application US/08464148  
Patent No. 5710026  
GENERAL INFORMATION:  
APPLICANT: Sprecher, Cindy A.  
TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,148  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/385,500  
FILING DATE: 08-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 13952-21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-148-4

Query Match 27.9%; Score 320; DB 1; Length 376;  
Best Local Similarity 46.3%; Pred. No. 2e-28;  
Matches 68; Conservative 28; Mismatches 45; Indels 6; Gaps 2;  
QY 47 PMSISSALAMVFMGAKGNTAAQMSQALCFKSGIGGEDGDIHRGFQSLLVAINRTDEYVLR 106  
DB 32 PVSISALAMVLLGAKGNTATQMAQALSLN---TEEDIHRAFQSLLVTEVKNKAGTQYLLR 87  
QY 107 TANGLFGEKSYDFTGTDSCGKEYQATIKOLDVNDTEKSTTRVNSWVADTK--KAWKI 164  
DB 88 TANRLFGKTCQFLSTFKESCLQFYHAEKLSFIRAAESRKHINTWVSKTEGKIEEL 147  
QY 165 IQTSLSHLEPFGIASSCYCKACLSQP 191  
DB 148 LPGSSIDAETRLVLVNAIYFKGKNWP 174

RESULT 6  
US-08-385-500-4  
Sequence 4, Application US/08385500  
Patent No. 5712117  
GENERAL INFORMATION:  
APPLICANT: Sprecher, Cindy A.  
TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND  
TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/385,500  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 13952-21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-385-500-4

Query Match 27.9%; Score 320; DB 1; Length 376;  
Best Local Similarity 46.3%; Pred. No. 2e-28;  
Matches 68; Conservative 28; Mismatches 45; Indels 6; Gaps 2;

QY 47 PMSISSALAWFMGKGNTRAAQMSQALCFKSGKIGGDIHRGFSQSLVAINRTDTEYVLR 106  
Db 32 PVSISALAWVLLGKAGNTATQMAQSLN-----TEEDIHRAFQSLTTEVNVKAGTQYLLR 87  
QY 107 TANGLFGEKSYDFTGFTDSCGKYQATIKQKDFVNDTEKSTTRVNSWADKT--KAWKI 164  
Db 88 TANLFGKCTQFLSTFKESCLQFYHAELKELSFIRAAEESRKHINTVWSKKTGKIEEL 147  
QY 165 IOTSLHLEEPGIASSSCYCKACLSQP 191  
Db 148 LPGSSIDAETRLVLVNAIFYKGNWP 174

RESULT 7  
US-08-846-784-4  
; Sequence 4, Application US/08846784  
; Patent No. 5747645  
; GENERAL INFORMATION:  
; APPLICANT: Sprecher, Cindy A.  
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND  
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/846,784  
; FILING DATE: 30-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/385,500  
; FILING DATE: 08-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.  
; REGISTRATION NUMBER: 31,990  
; REFERENCE/DOCKET NUMBER: 13952-21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 467-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 376 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-846-784-4

Query Match 27.9%; Score 320; DB 1; Length 376;  
Best Local Similarity 46.3%; Pred. No. 2e-28;  
Matches 68; Conservative 28; Mismatches 45; Indels 6; Gaps 2;  
QY 47 PMSISSALAWFMGKGNTRAAQMSQALCFKSGKIGGDIHRGFSQSLVAINRTDTEYVLR 106  
Db 32 PVSISALAWVLLGKAGNTATQMAQSLN-----TEEDIHRAFQSLTTEVNVKAGTQYLLR 87  
QY 107 TANGLFGEKSYDFTGFTDSCGKYQATIKQKDFVNDTEKSTTRVNSWADKT--KAWKI 164  
Db 88 TANLFGKCTQFLSTFKESCLQFYHAELKELSFIRAAEESRKHINTVWSKKTGKIEEL 147  
QY 165 IOTSLHLEEPGIASSSCYCKACLSQP 191  
Db 148 LPGSSIDAETRLVLVNAIFYKGNWP 174

RESULT 8  
US-08-568-147B-2  
; Sequence 2, Application US/08568147B  
; Patent No. 5783422  
; GENERAL INFORMATION:  
; APPLICANT: Suminami, Yoshinori  
; APPLICANT: Kato, Hiroshi  
; APPLICANT: Sekiguchi, Kiyoshi  
; APPLICANT: Takeda, Katsumichi  
; TITLE OF INVENTION: DNA FRAGMENT CODING FOR SQUAMOUS CELL  
; TITLE OF INVENTION: CARCINOMA-ASSOCIATED ANTIGEN  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/568,147B  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 800,952  
; FILING DATE: 02-DEC-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DiGiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 8425  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516-742-4343  
; TELEFAX: 516-742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 390 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-568-147B-2

Query Match 23.2%; Score 266; DB 1; Length 390;  
Best Local Similarity 37.9%; Pred. No. 3.4e-22;  
Matches 53; Conservative 28; Mismatches 47; Indels 12; Gaps 1;  
QY 32 PEKARGKQLKLIFFPMSISSALAWFMGKGNTRAAQMSQALCFKSGKIG----- 79  
Db 16 FOQFRKSKENNIFFYSPISITSLGCVLLGAKDNTAQQIKVLHFDQVNTTGAATVHV 75  
QY 80 GEDGDIHRGFSQSLVAINRTDTEYVLTANGLFGKESYDFTGFTDSCGKYQATIKOLD 139  
Db 76 DRSGNVHHQFKLLTEFNKSTDAYELKIANLFGKTYLFLQEYLDIAKKFYQTSVESVD 135  
QY 140 FVNDTEKSTTRVNSWADKT 159  
Db 136 FANAPESRRKINSWVESQT 155  
RESULT 9  
US-09-266-910-3  
; Sequence 3, Application US/09266910  
; Patent No. 634362  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Use of a recombinant protein as receptor of a  
; TITLE OF INVENTION: hepatitis virus  
; NUMBER OF SEQUENCES: 8



COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/266,910  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 390 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-09-266-910-3

Query Match 23.0%; Score 263; DB 4; Length 390;  
Best Local Similarity 37.9%; Pred. No. 7.6e-22;  
Matches 53; Conservative 27; Mismatches 48; Indels 12; Gaps 1;

QY 32 FEKARGKQLKQLIFFPMSTISSALAMVFMGAKGNTAAQMSQALCFESKIG----- 79  
DB 16 FQQRKSKENNIFYSPIISALGVLGAKDNTAQIKKVLHFDQVTTGKAATYHV 75  
QY 80 GEDGDIHRGFSQSLVAINRTDTEYVLTANGLFGEKSYDFLTGFTDSCGKFYQATIKOLD 139  
DB 76 DRSGNVHQQFQKLLTEFNKSTDAVELKIANKLFGEKTYLFQBYLDAIKKFYQTSVESVD 135  
QY 140 FVNDTEKSTTRVNSWVADKT 159  
DB 136 FANAPESPCKINSWVESQT 155

RESULT 10  
US-09-266-910-4  
Sequence 4, Application US/09266910  
Patent No. 6344362  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Use of a recombinant protein as receptor of a  
TITLE OF INVENTION: hepatitis virus  
NUMBER OF SEQUENCES: 8  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/266,910  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 390 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-09-266-910-4

Query Match 23.0%; Score 263; DB 4; Length 390;  
Best Local Similarity 37.9%; Pred. No. 7.6e-22;  
Matches 53; Conservative 27; Mismatches 48; Indels 12; Gaps 1;

QY 32 FEKARGKQLKQLIFFPMSTISSALAMVFMGAKGNTAAQMSQALCFESKIG----- 79  
DB 16 FQQRKSKENNIFYSPIISALGVLGAKDNTAQIKKVLHFDQVTTGKAATYHV 75  
QY 80 GEDGDIHRGFSQSLVAINRTDTEYVLTANGLFGEKSYDFLTGFTDSCGKFYQATIKOLD 139  
DB 76 DRSGNVHQQFQKLLTEFNKSTDAVELKIANKLFGEKTYLFQBYLDAIKKFYQTSVESVD 135  
QY 140 FVNDTEKSTTRVNSWVADKT 159  
DB 136 FANAPESPCKINSWVESQT 155

DB 76 DRSGNVHQQFQKLLTEFNKSTDAVELKIANKLFGEKTYLFQBYLDAIKKFYQTSVESVD 135  
QY 140 FVNDTEKSTTRVNSWVADKT 159  
DB 136 FANAPESPCKINSWVESQT 155

RESULT 11  
US-08-121-714-8  
Sequence 8, Application US/08121714  
Patent No. 5470970  
GENERAL INFORMATION:  
APPLICANT: Sager, Ruth  
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH  
TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,714  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER: 07/938,823  
FILING DATE: 09/01/92  
APPLICATION NUMBER: 07/844,296  
FILING DATE: 02/28/92  
APPLICATION NUMBER: 07/662,216  
FILING DATE: 02/28/91  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCES/DOCKET NUMBER: 00530/072001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-121-714-8

Query Match 22.4%; Score 257; DB 1; Length 375;  
Best Local Similarity 46.5%; Pred. No. 3.5e-21;  
Matches 53; Conservative 19; Mismatches 38; Indels 4; Gaps 1;

QY 47 PMSISSALAMVFMGAKGNTAAQMSQALCFESKIGGEDGDIHRGFSQSLVAINRTDTEYVLR 106  
DB 32 PLSSISSALAMIFLGTRGNTAAQVSKALYFDTV-----EDIHSRFQSLNADINKPGAPYILK 87  
QY 107 TANGLFGEKSYDFLTGFTDSCGKFYQATIKOLDVNDTEKSTTRVNSWVADKT 160  
DB 88 LANRLYGEKTYNFLADFLASTQKMYGASLASVDFOQAPEDARKEINWVKGOTE 141

RESULT 12  
US-08-477-108A-8  
Sequence 8, Application US/08477108A  
Patent No. 5801001  
GENERAL INFORMATION:

APPLICANT: Sager, Ruth  
APPLICANT: Zou, Zhiqiang  
APPLICANT: Anisowicz, Anthony  
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH  
TUMOR SUPPRESSING ACTIVITY  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,108A  
FILING DATE: June 7, 1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/121,714  
FILING DATE: 09/01/93  
APPLICATION NUMBER: 07/938,823  
FILING DATE: 09/01/92  
APPLICATION NUMBER: 07/844,296  
FILING DATE: 02/28/92  
APPLICATION NUMBER: 07/662,216  
FILING DATE: 02/28/91  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 06570/002002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-477-108A-8  
Query Match 22.4%; Score 257; DB 1; Length 375;  
Best Local Similarity 46.5%; Pred. No. 3.5e-21;  
Matches 53; Conservative 19; Mismatches 38; Indels 4; Gaps 1;  
QY 47 PMSISSALAMVFMGAKGNVTAQAQMSQALCFSGKIGGEDGDIHRGFSQSLVAINRTDTEYVLR 106  
Db 32 PLSSISALAMIFLGRGNTAAQVSKALYFDTV----EDIHSRFQSLNADINKPGAPYILK 87  
QY 107 TANGLFGEKSYDFLTGFTDSCGKFYQATIKQLDFVNDTEKSTTRVNSWVADTK 160  
Db 88 LANRLYGEKTYNFLADFLASTQKMYGAELASVDFOQAPEDARKEINWVKGQTE 141  
RESULT 13  
US-08-477-112-8  
Sequence 8, Application US/08477112  
Patent No. 5905023  
GENERAL INFORMATION:  
APPLICANT: Sager, Ruth  
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH  
TUMOR SUPPRESSING ACTIVITY  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston

STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,112  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/121,714  
FILING DATE: 09/01/93  
APPLICATION NUMBER: 07/938,823  
FILING DATE: 09/01/92  
APPLICATION NUMBER: 07/844,296  
FILING DATE: 02/28/92  
APPLICATION NUMBER: 07/662,216  
FILING DATE: 02/28/91  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 06570/002003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-477-112-8  
Query Match 22.4%; Score 257; DB 2; Length 375;  
Best Local Similarity 46.5%; Pred. No. 3.5e-21;  
Matches 53; Conservative 19; Mismatches 38; Indels 4; Gaps 1;  
QY 47 PMSISSALAMVFMGAKGNVTAQAQMSQALCFSGKIGGEDGDIHRGFSQSLVAINRTDTEYVLR 106  
Db 32 PLSSISALAMIFLGRGNTAAQVSKALYFDTV----EDIHSRFQSLNADINKPGAPYILK 87  
QY 107 TANGLFGEKSYDFLTGFTDSCGKFYQATIKQLDFVNDTEKSTTRVNSWVADTK 160  
Db 88 LANRLYGEKTYNFLADFLASTQKMYGAELASVDFOQAPEDARKEINWVKGQTE 141  
RESULT 14  
PCT-US93-08322-8  
Sequence 8, Application PC/TUS9308322  
GENERAL INFORMATION:  
APPLICANT: Sager, Ruth  
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08322  
FILING DATE:  
CLASSIFICATION:

RESULT 15  
US-07-768-286B-6  
; Sequence 6, Application US/07768286B  
; Patent No. 544153  
; GENERAL INFORMATION:  
; APPLICANT: GOSS, Neil H.  
; APPLICANT: RICHARDSON, Michael A.  
; TITLE OF INVENTION: VARIANTS OF PAI-2  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; City: Alexandria  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/768,286B  
; FILING DATE: 19911011  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/AU90/00603  
; FILING DATE: 20-DEC-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 16786/157 CHAC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 6:



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 10:26:02 ; Search time 56 Seconds  
(without alignments)  
380.752 Million cell updates/sec

Title: US-09-912-628-7  
Perfect score: 1145  
Sequence: 1 HELR5WAAARRTGARRHGCS.....SIPKCN5PVTPHGWNXPSSL 215

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

1 number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_5/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1143	99.8	215	9	US-09-912-628-7
2	1143	99.8	215	9	US-10-116-166-5
3	896	78.3	199	9	US-09-912-628-15
4	896	78.3	199	9	US-10-116-166-13
5	893	78.0	202	9	US-09-912-628-17
6	893	78.0	202	9	US-10-116-166-15
7	885.5	77.3	201	9	US-09-912-628-8
8	885.5	77.3	201	9	US-10-116-166-6
9	403.5	35.2	439	10	US-09-925-300-1440
10	392.5	34.3	374	9	US-09-974-298-95
11	278	24.3	390	1	US-08-731-566-2
12	240	21.0	437	10	US-09-925-301-1200
13	229	20.0	41	9	US-09-912-628-13
14	229	20.0	41	9	US-10-116-166-11
15	224	19.6	391	10	US-09-735-705-110
16	224	19.6	391	10	US-09-850-716A-110
17	234	19.6	391	10	US-09-897-778-110
18	223.5	19.5	415	10	US-09-902-684-4
19	222	19.4	42	9	US-09-912-628-11

20	222	19.4	42	9	US-10-116-166-9	Sequence 9, Appli
21	219.5	19.2	400	10	US-09-735-705-112	Sequence 112, App
22	219.5	19.2	400	10	US-09-850-716A-112	Sequence 112, App
23	219.5	19.2	400	10	US-09-897-778-112	Sequence 112, App
24	200	17.5	39	9	US-09-912-628-10	Sequence 10, Appli
25	200	17.5	39	9	US-10-116-166-8	Sequence 8, Appli
26	198	17.3	617	9	US-10-012-896-947	Sequence 947, App
27	198	17.3	617	9	US-09-895-814-947	Sequence 947, App
28	197	17.2	38	9	US-09-912-628-12	Sequence 12, Appli
29	197	17.2	38	9	US-10-116-166-10	Sequence 10, Appli
30	196	17.1	379	9	US-10-012-896-946	Sequence 946, App
31	196	17.1	379	9	US-09-895-814-946	Sequence 946, App
32	196	17.1	413	9	US-10-222-857-4	Sequence 4, Appli
33	186.5	16.3	418	9	US-10-135-629-2	Sequence 2, Appli
34	185.5	16.2	418	9	US-10-135-629-1	Sequence 1, Appli
35	175	15.3	380	9	US-10-091-442-34	Sequence 34, Appli
36	175	15.3	380	10	US-09-140-719-34	Sequence 34, Appli
37	174	15.2	315	1	US-08-731-566-4	Sequence 4, Appli
38	174	15.2	363	10	US-09-755-665-58	Sequence 58, Appli
39	173.5	15.2	377	9	US-10-165-605A-27	Sequence 27, Appli
40	173.5	15.2	377	10	US-09-910-430-27	Sequence 27, Appli
41	170	14.8	418	9	US-09-924-340-92	Sequence 92, Appli
42	170	14.8	418	9	US-09-992-600A-92	Sequence 92, Appli
43	170	14.8	418	9	US-10-000-489-92	Sequence 92, Appli
44	169	14.8	394	9	US-09-993-180-6	Sequence 6, Appli
45	169	14.8	394	9	US-10-025-514-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-912-628-7  
; Sequence 7, Application US/09912628  
; Patent No. US20020160491A1  
; GENERAL INFORMATION:  
; APPLICANT: NI et al.  
; TITLE OF INVENTION: Human Serpin Polynucleotides, Polypeptides, and Antibodies  
; FILE REFERENCE: PT001P2  
; CURRENT APPLICATION NUMBER: US/09/912,628  
; CURRENT FILING DATE: 2001-07-26  
; PRIOR APPLICATION NUMBER: PCT/US01/02484  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,769  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: PCT/US00/05082  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (211)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-912-628-7

Query Match 99.8%; Score 1143; DB 9; Length 215;  
Best Local Similarity 100.0%; Pred. No. 5.7e-114;  
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	HELRSWAAARRTGARRHGCSIRSKWHICIKPEKARGKOLKQLIFPMSISSALAMVFMG	60
DB	1	HELRSWAAARRTGARRHGCSIRSKWHICIKPEKARGKOLKQLIFPMSISSALAMVFMG	60
QY	61	AKGNTAAQMSQALCFSGKIGEDGDIHRGFQSLIVAINRTDTYVLTANGLFGEKSYDFL	120
DB	61	AKGNTAAQMSQALCFSGKIGEDGDIHRGFQSLIVAINRTDTYVLTANGLFGEKSYDFL	120
QY	121	TGFTDCSGFYQATIKQLDPVNDTEKSTRVNSWADKTKAMKIIOTSLSHLEPFIASS	180

	Query Match	78.3%;	Score 896;	DB 9;	Length 199;
	Best Local Similarity	99.8%;	Pred. No. 1.2e-87;		
	Matches 171;	Conservative	0;	Mismatches 2;	Indels 0;
					Gaps 0;
Qy	43	LIFFPMSSISALAMVFPMGAKNTAAQMSQALCFPSKIGGEDGDIHRGFOSLIVAINRTDTE	102		
Db	27	LIFFPMSSISALAMVFPMGAKNTAAQMSQALCFPSKIGGEDGDIHRGFOSLIVAINRTDTE	86		
Qy	103	YVLRTANGLFGKESYDFLTGTFDSCGFYQATIIKQLDFVNPDEKSTTRVNSWADKTKAW	162		

Db 87 YVLRNGLFGEKSYDFTLGTGFTDSCGKYQATIKQLDFVNDTEKSTTRVNSWADKTKAW 146  
QY 163 KIIQTSLSHLEPFGIASSCYKACLSQPLLHSHIPKCNKSPVTPHGMWPPSL 215  
Db 147 KIIQTSLSHLEPFGIASSCYKACLSQPLLHSHIPKCNKSPVTPHGMWPPSL 199

RESULT 5  
US-09-912-628-17  
; Sequence 17, Application US/09912628  
; Patent No. US20020160491A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: Human Serpin Polynucleotides, Polypeptides, and Antibodies  
; FILE REFERENCE: PT001P2  
; CURRENT APPLICATION NUMBER: US/09/912,628  
; CURRENT FILING DATE: 2001-07-26  
; PRIOR APPLICATION NUMBER: PCT/US01/02484  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,769  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: PCT/US00/05082  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 202  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: SITE  
; LOCATION: (31)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (32)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (33)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (198)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; US-09-912-628-17

Query Match 78.0%; Score 893; DB 9; Length 202;  
Best Local Similarity 93.0%; Pred. No. 2.5e-87;  
Matches 173; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 34 KARGKQLKLIFF---PMSISSALAMVFMGAKGNTAAQMSQALCFKSKIGGEDGDIHRGF 89  
Db 17 KLGGENNNLFFSXXXPMSSISSALAMVFMGAKGNTAAQMSQALCFKSKIGGEDGDIHRGF 76

QY 90 QSLLVAINRTDTEVLRNGLFGEKSYDFTLGTGFTDSCGKYQATIKQLDFVNDTEKSTT 149  
Db 77 QSLLVAINRTDTEVLRNGLFGEKSYDFTLGTGFTDSCGKYQATIKQLDFVNDTEKSTT 136

QY 150 RVNSWADKTKAWKIIQTSLSHLEPFGIASSCYKACLSQPLLHSHIPKCNKSPVTPHGM 209  
Db 137 RVNSWADKTKAWKIIQTSLSHLEPFGIASSCYKACLSQPLLHSHIPKCNKSPVTPHGM 196

QY 210 WXPSSL 215  
Db 197 WXPSSL 202

RESULT 6  
US-10-116-166-15  
; Sequence 15, Application US/10116166  
; Publication No. US20030040097A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Serpin Polynucleotides, Polypeptides, and Antibodies

; FILE REFERENCE: PT001P1  
; CURRENT APPLICATION NUMBER: US/10/116,166  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 09/641,721  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: PCT/US00/05092  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: 60/122,276  
; PRIOR FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 60/124,094  
; PRIOR FILING DATE: 1999-03-12  
; PRIOR APPLICATION NUMBER: 60/149,452  
; PRIOR FILING DATE: 1999-08-18  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 202  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: SITE  
; LOCATION: (31)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (32)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (33)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (198)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; US-10-116-166-15

Query Match 78.0%; Score 893; DB 9; Length 202;  
Best Local Similarity 93.0%; Pred. No. 2.5e-87;  
Matches 173; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 34 KARGKQLKLIFF---PMSISSALAMVFMGAKGNTAAQMSQALCFKSKIGGEDGDIHRGF 89  
Db 17 KLGGENNNLFFSXXXPMSSISSALAMVFMGAKGNTAAQMSQALCFKSKIGGEDGDIHRGF 76

QY 90 QSLLVAINRTDTEVLRNGLFGEKSYDFTLGTGFTDSCGKYQATIKQLDFVNDTEKSTT 149  
Db 77 QSLLVAINRTDTEVLRNGLFGEKSYDFTLGTGFTDSCGKYQATIKQLDFVNDTEKSTT 136

QY 150 RVNSWADKTKAWKIIQTSLSHLEPFGIASSCYKACLSQPLLHSHIPKCNKSPVTPHGM 209  
Db 137 RVNSWADKTKAWKIIQTSLSHLEPFGIASSCYKACLSQPLLHSHIPKCNKSPVTPHGM 196

QY 210 WXPSSL 215  
Db 197 WXPSSL 202

RESULT 7  
US-09-912-628-8  
; Sequence 8, Application US/09912628  
; Patent No. US20020160491A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: Human Serpin Polynucleotides, Polypeptides, and Antibodies  
; FILE REFERENCE: PT001P2  
; CURRENT APPLICATION NUMBER: US/09/912,628  
; CURRENT FILING DATE: 2001-07-26  
; PRIOR APPLICATION NUMBER: PCT/US01/02484  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,769  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: PCT/US00/05082  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patentin Ver. 2.0

```

US-10-116-166-6

Query Match          77.3%; Score 885.5; DB 9; Length 201;
Best Local Similarity 93.0%; Pred. No. 1.6e-86;
Matches 172; Conservative 2; Mismatches 8; Indels 3; Gaps

QY 34 KARGKQLKQLIFF--MSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFQ 90
DB 17 KKLGENNSNNLFFSPXMSIISALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFQ 76

QY 91 SLVAINRDTTEYVLRTANGLFGEKSYDFLTGFTDSCGKFYQATIKQLDFVNDTEKSTTR 150
DB 77 SLVAINRDTTEYVLRTANGLFGEKSYDFLTGFTDSCGKFYQATIKQLDFVNDTEKSTTR 136

QY 151 VNSWADKTKAWKIIQTSLSHLEEPGIASSCCYCKACLSQPLLVHSIPKCNSEPVTPHGMW 210
DB 137 VNSWADKTKAWKIIQTSLSHLEEPGIASSCCYCKACLSQPLLVHSIPKCNSEPVTPHGMW 196

QY 211 XPPSL 215
DB 197 XPPSL 201

RESULT 9
US-09-925-300-1440
; Sequence 1440, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1440
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1440

Query Match          35.2%; Score 403.5; DB 10; Length 459;
Best Local Similarity 51.5%; Pred. No. 1.3e-34;
Matches 85; Conservative 29; Mismatches 36; Indels 15; Gaps

QY 34 KARGKQLKQLIFF-PMSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFQSL 92
DB 100 KTLCKDKNKVFSPMSMSALAMVFMGAKGNTAAQMAQILSFNKS GG-GGDIHQGFQSL 150

QY 93 LVAINRDTTEYVLRTANGLFGEKSYDFLTGFTDSCGKFYQATIKQLDFVNDTEKSTTRVN 155
DB 159 LTYENKGTGYLLRWANRLFGEKSCDPLSFSSFRSCQKFYQAEWELDFISAVEKSRKHIN 211

QY 153 SWADKTKAWKIIQTSLSHLEEPGIASSCCYCKACLSQPLLVHSI 197
DB 219 TWVAEKTEG-----KIAELLSPGSVDP-----LTRLVLNNAV 250

RESULT 10
US-09-974-298-95
; Sequence 95, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04

```





LENGTH: 41  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (37)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-912-628-13

Query Match 20.0%; Score 229; DB 9; Length 41;  
Best Local Similarity 100.0%; Pred. No. 2e-17;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 PGIASSCYCKACLSQPLLVHISPKNSPVTPHGMWXPSSL 215  
DB 1 PGIASSCYCKACLSQPLLVHISPKNSPVTPHGMWXPSSL 41

RESULT 14

US-10-116-166-11  
Sequence 11, Application US/10116166  
Publication No. US20030040097A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: Serpin Polynucleotides, Polypeptides, and Antibodies  
FILE REFERENCE: PT001P1  
CURRENT APPLICATION NUMBER: US/10/116,166  
CURRENT FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: 09/641,721  
PRIOR FILING DATE: 2000-08-21  
PRIOR APPLICATION NUMBER: PCT/US00/05092  
PRIOR FILING DATE: 2000-02-29  
PRIOR APPLICATION NUMBER: 60/122,276  
PRIOR FILING DATE: 1999-03-01  
PRIOR APPLICATION NUMBER: 60/124,094  
PRIOR FILING DATE: 1999-03-12  
PRIOR APPLICATION NUMBER: 60/149,452  
PRIOR FILING DATE: 1999-08-18  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 11  
LENGTH: 41  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (37)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-116-166-11

Query Match 20.0%; Score 229; DB 9; Length 41;  
Best Local Similarity 100.0%; Pred. No. 2e-17;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 PGIASSCYCKACLSQPLLVHISPKNSPVTPHGMWXPSSL 215  
DB 1 PGIASSCYCKACLSQPLLVHISPKNSPVTPHGMWXPSSL 41

RESULT 15

US-09-735-705-110  
Sequence 110, Application US/09735705  
Patent No. US2002052329A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Fan, Liqun  
APPLICANT: Kalos, Michael D.  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Hosken, Nancy  
APPLICANT: Fanger, Gary R.  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yabir A.W.

APPLICANT: Henderson, Robert A.  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Fanger, Neil  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.455C14  
CURRENT APPLICATION NUMBER: US/09/735,705  
CURRENT FILING DATE: 2000-12-12  
NUMBER OF SEQ ID NOS: 419  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 110  
LENGTH: 391  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-735-705-110

Query Match 19.6%; Score 224; DB 10; Length 391;  
Best Local Similarity 32.7%; Pred. No. 1.5e-15;  
Matches 53; Conservative 31; Mismatches 54; Indels 24; Gaps 4;

QY 38 KQLKQ-----LIFPPMSISSALAMVFMGAKGNTAAQMSQAL-----CFSKIGGEDGD-- 84  
DB 17 KELKTNDSGNIFSPVIGILLTAIGWLLGTRGATASQLEEVFHSSEKTKSSRIKAEKEVI 76  
QY 85 -----IHRGFQSLVAINRTDTTEYVLRNTANGLFGEKSYDELFTDSCGKFYQATIKQLD 139  
DB 77 ENTEAVHQPFQKELTEISKLTNDYELNITNRLFGEKTYLEQLKYLDYVEKYHASLEPVD 136  
QY 140 FVNDTEKSTTRVNSWVADTKAWKIIQTSLSHLEEPGIASSS 181  
DB 137 FVNADESRRKINSWVESKT-----NEKIDLPDGSISSS 172

Search completed: May 20, 2003, 10:35:38  
Job time : 57 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 10:22:17 ; Search time 44 Seconds  
(without alignments)  
469.748 Million cell updates/sec

Title: US-09-912-628-7  
Perfect score: 1145  
Sequence: 1 HELRSWAARRTGARRHGS .....SIPKCNPTVPHGMWPPSL 215  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	404.5	35.3	376	1 A48681	placental thrombin
2	392.5	34.3	374	2 A59273	proteinase inhibit
3	370.5	32.4	378	2 A57488	proteinase inhibit
4	320	27.9	376	2 B59273	proteinase inhibit
5	272	23.8	330	2 I38202	leupin precursor -
6	266	23.2	330	2 I38201	squamous cell carc
7	257	22.4	379	2 A42421	leukocyte elastase
8	255.5	22.3	378	2 S38962	serpin - pig
9	240	21.0	379	2 S27383	elastase inhibitor
10	224	19.6	391	2 JC7118	headpin serine pro
11	221.5	19.3	415	2 A32853	plasminogen activa
12	216.5	18.9	415	2 S20047	plasminogen activa
13	213.5	18.6	397	2 I39184	bomapsin - human
14	213	18.6	416	2 S19896	plasminogen activa
15	198.5	17.3	388	1 DYCH	ovalbumin-related
16	191.5	16.7	402	2 I49471	alpha-1 proteinase
17	190.5	16.6	413	2 I49452	alpha-1-antitryps
18	190.5	16.6	413	2 I49473	alpha-1 proteinase
19	189.5	16.6	413	2 I49474	alpha-1 proteinase
20	188.5	16.5	418	2 S23675	contrapsin-related
21	186.5	16.3	413	2 I49470	alpha-1 proteinase
22	186.5	16.3	413	2 I49472	alpha-1 proteinase
23	186.5	16.3	418	2 JH0494	alpha-1 antichymot
24	185	16.2	405	2 A33088	alpha-1-antitryps
25	185	16.2	405	2 A33088	alpha-1-antitryps
26	184.5	16.1	418	2 JX0129	contrapsin precurs
27	184	16.1	418	2 A53120	intracellular coag
28	183.5	16.0	405	2 A28321	corticosteroid-bin
29	183.5	16.0	413	2 I56481	alpha-1-proteinase

30	183	16.0	410	2 C39088	contrapsin precurs
31	182	15.9	388	2 B39088	alpha-1-antitryps
32	181.5	15.9	411	1 ITRT	alpha-1-antitryps
33	181.5	15.9	417	2 S19724	kalikrein-binding
34	181	15.8	416	2 S21097	alpha-1-antitryps
35	179	15.6	413	2 A54968	alpha-1-antitryps
36	178.5	15.6	383	2 A36117	corticosteroid-bin
37	175	15.3	409	1 ITBA	alpha-1-antitryps
38	175	15.3	416	1 ITSH	alpha-1-antitryps
39	174	15.2	413	2 JX0267	alpha-1-antitryps
40	174	15.2	413	2 S54981	alpha-1-antitryps
41	173.5	15.2	408	2 S11320	serine proteinase
42	173.5	15.2	436	2 A42440	estrogen-regulated
43	173	15.1	413	2 JX0154	alpha-1-antitryps
44	170.5	14.9	418	1 S31507	serine proteinase
45	169	14.8	418	1 ITHU	alpha-1-antitryps

ALIGNMENTS

RESULT 1

A48681  
placental thrombin inhibitor - human  
N; Alternate names: cytoplasmic antiproteinase; intracellular serine proteinase inhibitor.  
C: Species: Homo sapiens (man)  
C: Date: 07-Apr-1994 #sequence revision 07-Jul-1995 #text change 18-Jun-1999  
C: Accession: A48681; A54352; A46672; B46672; C46672; S35750  
R: Coughlin, P.; Sun, J.; Cerruti, L.; Salem, H.H.; Bird, P.  
Proc. Natl. Acad. Sci. U.S.A. 90, 9417-9421, 1993  
A: Title: Cloning and molecular characterization of a human intracellular serine protein.  
A: Reference number: A48681; MUID: 94022386; PMID: 8415716  
A: Accession: A48681  
A: Molecule type: mRNA  
A: Residues: 1-376 <COU>  
A: Cross-references: GB: Z22658; NID: 9297411; PIDN: CAA80373.1; PID: g297412  
A: Experimental source: Placenta  
A: Note: authors translated the codon CAA for residue 198 as Gly  
R: Morgenstern, K.A.; Sprecher, C.; Holch, L.; Foster, D.; Grant, F.J.; Ching, A.; Kisse  
Biochemistry 33, 3432-3441, 1994  
A: Title: Complementary DNA cloning and kinetic characterization of a novel intracellular  
A: Reference number: A54352; MUID: 94183847; PMID: 8136380  
A: Accession: A54352  
A: Molecule type: mRNA  
A: Residues: 1-174, 'E', 176-361, 'S', 363-376 <MOR>  
A: Cross-references: GB: S69272; NID: g546087; PIDN: AAB30320.1; PID: g546088  
A: Experimental source: Placenta  
A: Note: sequence extracted from NCBI backbone (NCBIN: 145231, NCBIP: 145232)  
R: Coughlin, P.B.; Tetaz, T.; Salem, H.H.  
J. Biol. Chem. 268, 9541-9547, 1993  
A: Title: Identification and purification of a novel serine proteinase inhibitor.  
A: Reference number: A46672; MUID: 93252826; PMID: 8486644  
A: Accession: A46672  
A: Molecule type: protein  
A: Residues: 47-60; 63-81; 91-98 <CO2>  
A: Experimental source: Placenta, leukemic cell line K562  
A: Note: sequence modified after extraction from NCBI backbone  
C: Genetics:  
A: Gene: GDB: PI6  
A: Cross-references: GDB: 252025; OMIM: 173321  
A: Map position: 6p25-6p24.3  
A: Superfamily: antithrombin III  
C: Keywords: blocked amino end; cytosol; serine proteinase inhibitor  
F: 341/Inhibitory site: Arg (thrombin) #status predicted

Query Match 35.3%; Score 404.5; DB 1; Length 376;  
Best Local Similarity 51.5%; Pred. No. 3.7e-30;

Matches 85; Conservative 29; Mismatches 36; Indels 15; Gaps 4;

34 KARGKQLKQLIFF-PWSISSALAVFMGKNTAQMSSQALCFKIGGEDGDIHRGFQSL 92

DB 17 KTLGKDNKSNVFPSPMWSKALAVYMGKNTAQMSSQALCFKIGGEDGDIHRGFQSL 75

A:Accession: P13437  
A:Status: preliminary  
A:Map position: 13  
A:Molecule type: mRNA  
A:Residues: 1-378 <SUN>  
A:Cross-references: GB:U25844; NID:g818902; PIDN:AAA79684.1; PID:g818903  
C:Genetics:  
A:Gene: Spi3  
A:Superfamily: antithrombin III  
C:Keywords: serine proteinase inhibitor  
P.343/Inhibitory site: Arg (unidentified proteinase) #status predicted

RESULT 5  
I38202  
leupin precursor - human  
N;Alternate names: proteinase inhibitor 11 (Pill); squamous cell carcinoma  
C;Species: Homo sapiens (man)  
C;Date: 23-Feb-1996 #sequence revision 23-Feb-1996 #text\_change 26-May-2000  
C;Accession: I38202; S66675; S57522  
R;Schneider, S.S.; Schick, C.; Fish, K.E.; Miller, E.; Pena, J.C.; Treter,  
Proc. Natl. Acad. Sci. U.S.A. 92: 3147-3151, 1995  
A;Title: A serine proteinase inhibitor locus at 19q21.3 contains a tandem  
A;Reference number: I38200; MUID:95241462; PMID:7724531  
A;Accession: I38202  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA

A;Residues: 1-390 <SCH>  
A;Cross-references: EMBL:U19576; GB:U19569; NID:g852466  
R;Barnes, R.C.; Worrall, D.M.  
FEBS Lett. 373, 61-65, 1995  
A;Title: Identification of a novel human serpin gene; cloning sequencing and expression  
A;Reference number: S66675; MUID:96013887; PMID:7589435  
A;Accession: S66675  
A;Molecule type: mRNA  
A;Gene: GDB:SCCA2; P111  
A;Cross-references: EMBL:X89015; NID:g887464; PIDN:CAA61420.1; PID:g887465  
C;Genetics:  
A;Map position: 18q21.3-18q21.3  
A;Introns: 55/3; 74/3; 117/3; 157/1; 204/3; 256/3  
C;Superfamily: antithrombin III  
C;Keywords: glycoprotein; serine proteinase inhibitor  
F;1-22/Domain: signal sequence status predicted <SG>  
F;23-384/Product: Leupin #status predicted <MAR>  
F;65-93,170,376/Binding site: carbohydrate (Asn) #status predicted  
F;34/Inhibitory site: Leu (unidentified proteinase) #status predicted

Query Match 23.8%; Score 272; DB 2; Length 390;  
Best Local Similarity 38.6%; Pred. No. 1e-17;  
Matches 54; Conservative 27; Mismatches 47; Indels 12; Gaps 1;

QY 32 FEKARGKQLKQLFFPMSSISALAMVFMGAKNTAAQMSQALCFKSGKIG----- 79  
DB 16 FQPRKSKENNIFYPSPISITSLGMLVLLGAKNTAAQIKKVLHFDQVTENTTEKAATYHV 75  
QY 80 GEDGDIHRGFQSLVAINRTDTEVLRTANGLFGEKSYDFLTGFTDSCGKFGYQATIKOLD 139  
DB 76 DRSGNVHGFQKLLTFEKNKTDAYELKIANKLFGEKTYLFLQELDAIKKFGYQTSVESTD 135  
QY 140 FVNDTEKSTRVNSWVADKT 159  
DB 136 FANAPEESRKINSWVESQT 155

RESULT 6  
I38201  
squamous cell carcinoma antigen 1 - human  
C;Species: Homo sapiens (man)  
C;Date: 23-Feb-1996 #sequence revision 23-Feb-1996 #text\_change 26-May-2000  
C;Accession: I38201; I38200; G01631; J00967  
R;Schneider, S.S.; Schick, C.; Fish, K.E.; Miller, E.; Pena, J.C.; Treter, S.D.; Hui, S.  
Proc. Natl. Acad. Sci. U.S.A. 92, 3147-3151, 1995  
A;Title: A serine proteinase inhibitor locus at 18q21.3 contains a tandem duplication of  
A;Reference number: I38200; MUID:95241462; PMID:7724531  
A;Accession: I38201  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-390 <SCH1>  
A;Cross-references: EMBL:U19568; GB:U19558; NID:g1172085  
A;Accession: I38200  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-117 <SCH2>  
A;Cross-references: EMBL:U19562; NID:g897835; PIDN:AAA86316.1; PID:g897844  
R;Silverman, G.A.  
submitted to the EMBL Data Library, January 1995  
A;Reference number: G07968  
A;Accession: G01631  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-350, 'G', 352-390 <SLU>  
A;Cross-references: EMBL:U19556; NID:g1276435; PID:g1052869  
R;Suminami, Y.; Kishi, F.; Sekiguchi, K.; Kato, H.  
Biochem. Biophys. Res. Commun. 181, 51-58, 1991  
A;Title: Squamous cell carcinoma antigen is a new member of the serine protease inhibitor  
A;Reference number: J00966; MUID:92068241; PMID:1958219  
A;Accession: J00966  
A;Molecule type: mRNA

A;Residues: 1-350, 'G', 352-356, 'A', 358-390 <SUM1>  
A;Cross-references: GB:S66896; NID:g239551; PIDN:AAB20405.1; PID:g239552  
A;Accession: J00967  
A;Molecule type: protein  
A;Residues: 11-21,231-237;240-256;303-325 <SUM2>  
C;Comment: This antigen probably acts as a proteinase inhibitor to modulate the host i  
C;Genetics:  
A;Gene: GDB:SCAL; SCC  
A;Cross-references: GDB:625364; OMIM:600517  
A;Map position: 18q21.3-18q21.3  
A;Introns: 55/3; 74/3; 117/3; 157/1; 204/3; 256/3  
C;Superfamily: antithrombin III  
C;Keywords: cysteine proteinase inhibitor; glycoprotein  
F;65-93,171,376/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;34/Inhibitory site: Ser (cathepsin L) #status predicted

Query Match 23.2%; Score 266; DB 2; Length 390;  
Best Local Similarity 37.9%; Pred. No. 3.8e-17;  
Matches 53; Conservative 28; Mismatches 47; Indels 12; Gaps 1;

QY 32 FEKARGKQLKQLFFPMSSISALAMVFMGAKNTAAQMSQALCFKSGKIG----- 79  
DB 16 FQPRKSKENNIFYPSPISITSLGMLVLLGAKNTAAQIKKVLHFDQVTENTTEKAATYHV 75  
QY 80 GEDGDIHRGFQSLVAINRTDTEVLRTANGLFGEKSYDFLTGFTDSCGKFGYQATIKOLD 139  
DB 76 DRSGNVHGFQKLLTFEKNKTDAYELKIANKLFGEKTYLFLQELDAIKKFGYQTSVESVD 135  
QY 140 FVNDTEKSTRVNSWVADKT 159  
DB 136 FANAPEESRKINSWVESQT 155

RESULT 7  
A42421  
leukocyte elastase inhibitor - horse  
N;Alternate names: plasminogen activator inhibitor-2 homolog  
C;Species: Equus caballus (domestic horse)  
C;Date: 04-Mar-1993 #sequence revision 18-Nov-1994 #text\_change 02-Jun-2000  
C;Accession: A42421; A37276; S34062  
R;Dubin, A.; Travis, J.; Enghild, J.J.; Potempa, J.  
J. Biol. Chem. 267, 6576-6583, 1992  
A;Title: Equine leukocyte elastase inhibitor. Primary structure and identification as a  
A;Reference number: A42421; MUID:92202200; PMID:1551869  
A;Accession: A42421  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-379 <DUB1>  
A;Cross-references: PIDN:AAB21885.1; PID:g247842  
A;Experimental source: leukocyte  
A;Note: sequence extracted from NCBI backbone (NCBI:89849)  
R;Dubin, A.; Travis, J.; Enghild, J.J.; Potempa, J.  
submitted to the Protein Sequence Database, December 1991  
A;Reference number: A37276  
A;Accession: A37276  
A;Molecule type: protein  
A;Residues: 1-41, 'E', 43-325, 'VD', 326-379 <DUB2>  
R;Kordula, T.; Dubin, A.; Schooltink, H.; Koj, A.; Heinrich, P.C.; Rose-John, S.  
Biochem. J. 293, 187-193, 1993  
A;Title: Molecular cloning and expression of an intracellular serpin: an elastase inhib  
A;Reference number: S34062; MUID:93319507; PMID:7687128  
A;Accession: S34062  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-379 <KOR>  
A;Cross-references: GB:M91161; NID:g164240; PIDN:AAA97513.1; PID:g164241  
C;Superfamily: antithrombin III  
C;Keywords: serine proteinase inhibitor

Query Match 22.4%; Score 257; DB 2; Length 379;  
Best Local Similarity 46.5%; Pred. No. 2.6e-16;  
Matches 53; Conservative 19; Mismatches 38; Indels 4; Gaps 1;

Query Match 21.0%; Score 240; DB 2; Length 379;  
Best Local Similarity 43.0%; Pred. No. 1e-14;  
Matches 49; Conservative 23; Mismatches 38; Indels 4; Gaps 1;  
QY 47 PMSISSALAMVFMGAKGNTAAQMSQALCFKIGGEDGDIHRGFSQSLLVAINRTDTEYVLR 106  
DB 32 PLSISSALAMIFLGRGNTAAQVSKALYFDTV---EDIHRSFQSLNADINKPGAPYILK 87  
QY 107 TANGLFGEKSYDFLTGTDCGKGYQATIKQLDFVNDTEKSTTRVNSWVADTK 160  
DB 88 LANRLYGEKTYNFLADFLASTQKMGAEALASVDFOQAPEDARKNEINWVGQTE 141  
RESULT 8  
S38962  
serpin - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 02-Jul-1998  
C:Accession: S38962  
R:Teuchauer, W.F.; Mentelle, R.; Sommerhoff, C.P.  
Eur. J. Biochem. 217, 519-526, 1993  
A:Title: Primary structure of a porcine leukocyte serpin.  
A:Reference number: S38962; MUID:94039085; PMID:7901009  
A:Accession: S38962  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-378 <TES>  
A:Note: the sequence from Fig. 6 is inconsistent with that from Fig. 5 in having 256-Asn  
C:Superfamily: antithrombin III

Query Match 22.3%; Score 255.5; DB 2; Length 378;  
Best Local Similarity 39.1%; Pred. No. 3.6e-16;  
Matches 59; Conservative 27; Mismatches 48; Indels 17; Gaps 3;  
QY 47 PMSISSALAMVFMGAKGNTAAQMSQALCFKIGGEDGDIHRGFSQSLLVAINRTDTEYVLR 106  
DB 32 PFSISSALAMILLGRGNTAAQMSKALHFDTV---KDIHRSFQSLNADINKGCASYILK 87  
QY 107 TANGLFGEKSYDFLTGTDCGKGYQATIKQLDFVNDTEKSTTRVNSWVADTKAWKIIQ 166  
DB 88 LANRLYGEKTYHFLPEFLASTQKTYGAELASVDPLRASEEARKAINWVKQTEG----- 142  
QY 167 TSLSHLEBPGIASSCYKACLSQBLVHSI 197  
DB 143 -KIPPELLASGVVDSA-----TKLVLVNAI 165

RESULT 9  
S27383  
elastase inhibitor - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 05-Sep-1997  
C:Accession: S27383; S65750  
A:Title: Sequence and molecular characterization of human monocyte/neutrophil elastase  
A:Reference number: S27383; MUID:92302296; PMID:1376927  
A:Accession: S27383  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-379 <REM>  
A:Cross-references: EMBL:M93056  
R:Packard, B.Z.; Lee, S.S.; Remold-O'Donnell, E.; Komoriya, A.  
Biochim. Biophys. Acta 1269, 41-50, 1995  
A:Title: A serpin from human tumor cells with direct lymphoid immunomodulatory activity.  
A:Reference number: S65750; MUID:96049524; PMID:7578269  
A:Accession: S65750  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 57-69; 97-110; 111-129; 204-213; 216-244; 255-271, 'X', 273-274; 291-301 <PAC>  
C:Genetics:  
A:Gene: GDB:ELANH2; EI; P12  
A:Cross-references: GDB:132914; OMIM:130135  
A:Map position: 6p25-6p24.3  
C:Superfamily: antithrombin III  
C:Keywords: serine proteinase inhibitor

Query Match 21.0%; Score 240; DB 2; Length 379;  
Best Local Similarity 43.0%; Pred. No. 1e-14;  
Matches 49; Conservative 23; Mismatches 38; Indels 4; Gaps 1;  
QY 47 PMSISSALAMVFMGAKGNTAAQMSQALCFKIGGEDGDIHRGFSQSLLVAINRTDTEYVLR 106  
DB 32 PFSISSALAMVFLGRGNTAAQLSKTFHTV---EEVHRSFQSLNADINKRGASYILK 87  
QY 107 TANGLFGEKSYDFLTGTDCGKGYQATIKQLDFVNDTEKSTTRVNSWVADTK 160  
DB 88 LANRLYGEKTYNFLPEFLVSTQKTYGADLASVDFOHASEDARKTINQWVGQTE 141  
RESULT 10  
JC7118  
headpin serine proteinase inhibitor - human  
C:Species: Homo sapiens (man)  
C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 26-May-2000  
C:Accession: JC7118  
R:Spring, P.; Nakashima, T.; Frederick, M.; Henderson, Y.; Clayman, G.  
Biochem. Biophys. Res. Commun. 264, 299-304, 1999  
A:Title: Identification and cDNA cloning of headpin, a novel differentially expressed s  
A:Reference number: JC7118; MUID:99458661; PMID:10527881  
A:Accession: JC7118  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-391 <SPR>  
A:Cross-references: GB:AF169949; NID:G5911368; PIDN:AAD55765.1; PID:G5911369  
C:Genetics:  
A:Map position: 18q21.3-19q22  
C:Superfamily: antithrombin III  
C:Keywords: serine proteinase inhibitor  
P:355/Inhibitory site: Val (unidentified proteinase) #status predicted  
Query Match 19.6%; Score 224; DB 2; Length 391;  
Best Local Similarity 32.7%; Pred. No. 3.4e-13;  
Matches 53; Conservative 31; Mismatches 54; Indels 24; Gaps 4;  
QY 38 KOLKQ-----LIFFPMSISSALAMVFMGAKGNTAAQMSQAL-----CFSKIGGEDG-- 84  
DB 17 KELKNTDGNIFPSPVIGLITAGVLLGTRGATASQLEEVFHSSEKETSRIKAEKEVI 76  
QY 85 -----IHRGFSQSLLVAINRTDTEYVLTANGLFGEKSYDFLTGTDCGKGYQATIKQLD 139  
DB 77 ENTEAVHQQFKFLTEISKLTNDYELNITNRLGFKYTLFLQKLYVVEKYYHASLEPVD 136  
QY 140 FVNDTEKSTTRVNSWVADTKAWKIIQTSLSHLEEPGIASSS 181  
DB 137 FVNADSRKINSWVESKT-----NEKIKLFPDGSISSS 172  
RESULT 11  
A32853  
plasminogen activator inhibitor 2 precursor - human  
N:Alternate names: urokinase inhibitor  
C:Species: Homo sapiens (man)  
C:Date: 22-Nov-1989 #sequence\_revision 22-Nov-1989 #text\_change 20-Oct-2000  
C:Accession: A32853; I54218; A31366; A46543; A29362; A53815; A25021; S02435; I5.  
R:Ye, R.D.; Ahern, S.M.; Le Beau, M.M.; Lebo, R.V.; Sadler, J.E.  
J. Biol. Chem. 264, 5495-5502, 1989  
A:Title: Structure of the gene for human plasminogen activator inhibitor-2. The nearest  
A:Reference number: A32853; MUID:89174589; PMID:2494165  
A:Accession: A32853  
A:Molecule type: DNA  
A:Residues: 1-415 <YER>  
A:Cross-references: GB:M24651; GB:M24652; GB:M24653; GB:M24654; GB:M24655; GB:M24656; I5.  
R:Samia, J.A.; Alexander, S.J.; Horton, K.W.; Aaron, P.E.; Byers, M.G.; Shows, T.B.  
Genomics 6, 159-167, 1990  
A:Title: Chromosomal organization and localization of the human urokinase inhibitor gen  
A:Reference number: I54218; MUID:90152678; PMID:2303256  
A:Accession: I54218  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA



Nucleic Acids Res. 10, 4363-4382, 1982  
A:Title: The ovalbumin gene family: complete sequence and structure of the Y gene.  
A:Reference number: A01244; MUID:63014329; PMID:7122240  
A:Accession: A01244  
A:Molecule type: DNA  
A:Residues: 1-388 <HEI>  
A:Cross-references: GB:J00922; GB:V00439; NID:g212899; PIDN:AAA68882.1; PID:g212900  
C:Genetics:  
A:Introns: 56/3; 73/3; 116/3; 156/1; 203/3; 255/3  
C:Superfamily: antithrombin III  
C:Keywords: glycoprotein; phosphoprotein; serine proteinase inhibitor  
F:74-121/Disulfide bonds: #status predicted  
F:293/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:345/Binding site: phosphate (Ser) (covalent) #status predicted  
Query Match 17.3%; Score 198.5; DB 1; Length 389;  
Best Local Similarity 30.1%; Pred. No. 8.3e-11;  
Matches 43; Conservative 32; Mismatches 57; Indels 11; Gaps 2;  
QY 28 CIKPEKARGKQLKQ-LIFFPMSISSALAMVFMGAKGNTAAQMSQALCFKSGEDGD-- 84  
Db 12 CFDEVENEMKVHVNNENILYCPILITLALAMVYLGARGNTESQMKVLFHDSITGAGSTTD 71  
QY 85 -----IHRGFQSLVAINRTDTEYVLRVLTANGLFGEKSYDFLTGTDSCGKFFQATIK 136  
Db 72 SQCGSEYVHNLFKELLSEITRPNTATYSLIADKLYVDKTFVLPYLSLSCARKFYTGVE 131  
QY 137 QLDVNDTEKSTTRVNSWADKT 159  
Db 132 EYVNTAAAEARQLINSWVKET 154

Search completed: May 20, 2003, 10:26:44  
Job time : 46 secs

R:Riewald, M.; Schleef, R.R.  
J. Biol. Chem. 270, 26754-26757, 1995  
A:Title: Molecular cloning of bomapin (potease inhibitor 10), a novel human serpin that  
A:Reference number: I39184; MUID:96070759; PMID:7592909  
A:Accession: I39184  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-397 <RES>  
A:Cross-references: EMBL:U35459; NID:g1065408; PIDN:AAC50282.1; PID:g1065409  
C:Genetics:  
A:Gene: GDB:P110  
A:Cross-references: GDB:636283  
A:Map position: 14q32.1-14q32.1  
C:Superfamily: antithrombin III  
C:Keywords: serine proteinase inhibitor  
Query Match 18.6%; Score 213.5; DB 2; Length 397;  
Best Local Similarity 32.2%; Pred. No. 3.3e-12;  
Matches 47; Conservative 34; Mismatches 44; Indels 21; Gaps 2;  
QY 33 EKARGKQLKQLIFFPMSISSALAMVFMGAKGNTAAQMSQALCFKSGIGG----- 80  
Db 21 ESAQG---KNIPFSSISITSLTIVLGAQTAAQMAQVLOFNDRDQGVKCDPESEKKRK 77  
QY 81 -----EDGDIHRGFQSLVAINRTDTEYVLRVLTANGLFGEKSYDFLTGTDSCGKFFQAT 134  
Db 78 MEFNLNSEEIHSDFQTLSEILKPNDDVLLKTANAIYGEKTYAFHNKYLEDMKTYFGAE 137  
QY 135 IKOLDVNDTEKSTTRVNSWADKT 160  
Db 138 PQPVNFVEASDQIRKINSWVERQTE 163

RESULT 14  
S19896  
plasminogen activator inhibitor 2 type A - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jul-1999  
C:Accession: S19896  
R:Grundmann, U.; Rein, T.  
A:Description: Complete cDNA sequence encoding plasminogen activator inhibitor 2 type A  
A:Reference number: S19896  
A:Molecule type: mRNA  
A:Residues: 1-416 <GRU>  
A:Cross-references: EMBL:X64563; NID:g56834; PIDN:CAA45864.1; PID:g56835  
C:Superfamily: antithrombin III

Query Match 18.6%; Score 213; DB 2; Length 416;  
Best Local Similarity 31.9%; Pred. No. 3.9e-12;  
Matches 53; Conservative 22; Mismatches 55; Indels 36; Gaps 3;  
QY 29 IKPPEKARGKQLKQLIFFPMSISSALAMVFMGAKGNTAAQMSQALCFKSGIGG----- 82  
Db 16 LKQIEQSNSTQ--NIFISPSNISSTLAIFLGAQNTAEQMAKVLNPDKIGSYDLTPGNP 73  
QY 83 -----GDHIGFQSLVAINRTDT--EYVLRVLTANGLFGE 114  
Db 74 ENFHGCDFAHQIRDNYFVAILQAQARDKIHSAFSSLSSTINTPRIGDYLLLESANKLFGE 133  
QY 115 KSYDPLAGTDSGKFFQATIKOLDVNDTEKSTTRVNSWADKT 160  
Db 134 KSARFKEEYIQRCKKYSTEPEAVDFLECANEARAKKINSWVKQTQK 179  
RESULT 15  
DYCH  
ovalbumin-related Y protein - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 18-Jun-1999  
C:Accession: A01244  
R:Heilig, R.; Muraskowsky, R.; Kloefer, C.; Mandel, J.L.



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 10:19:11 ; Search time 13 Seconds  
(without alignments)  
685.955 Million cell updates/sec

Title: US-09-912-628-7

Perfect score: 1145

Sequence: 1 HELRSWAARRTGARRHGS.....SIPKCNPTVPHGMXPPSL 215

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Spaced: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	404.5	35.3	376	1 PT16 HUMAN	P35237 homo sapien
2	392.5	34.3	374	1 SPB8 HUMAN	P50452 homo sapien
3	388	33.9	378	1 PT16 BOVIN	O02739 bos taurus
4	370.5	32.4	378	1 PT16 MOUSE	Q60854 mus musculus
5	320	27.9	376	1 SPB9 HUMAN	P50453 homo sapien
6	272	23.8	390	1 SCC2 HUMAN	P48594 homo sapien
7	266	23.2	390	1 SCC1 HUMAN	P29508 homo sapien
8	257	22.4	379	1 ILEU HORSE	P05619 equus caball
9	255.5	22.3	378	1 ILEU PIG	P80229 sus scrofa
10	240	21.0	379	1 ILEU HUMAN	P30740 homo sapien
11	224	19.6	391	1 SB13 HUMAN	Q9ui18 homo sapien
12	221.5	18.3	415	1 PA12 HUMAN	P05120 homo sapien
13	216.5	18.9	415	1 PA12 MOUSE	P12388 mus musculus
14	213.5	18.6	397	1 SB10 HUMAN	P48595 homo sapien
15	213	18.6	416	1 PA12 RAT	P29524 rattus norv
16	207	18.1	413	1 ALST TAMSI	O54760 tamias sibi
17	202	17.6	413	1 ALST TAMSI	O54759 tamias sibi
18	198.5	17.3	388	1 OVAY CHICK	O54757 tamias sibi
19	197	17.2	413	1 ALUM TAMSI	Q96p15 homo sapien
20	196	17.1	392	1 SB11 HUMAN	O90955 tamias sibi
21	196	17.1	413	1 HP55 TAMSI	Q96p63 homo sapien
22	193.5	16.9	405	1 SB12 HUMAN	O54758 tamias sibi
23	193	16.9	413	1 ALMS TAMSI	P22599 mus musculus
24	190.5	16.6	413	1 ALT2 MOUSE	Q00897 mus musculus
25	190.5	16.6	413	1 ALT4 MOUSE	Q00898 mus musculus
26	189.5	16.6	413	1 ALT5 MOUSE	P07758 mus musculus
27	186.5	16.3	413	1 ALT1 MOUSE	Q00896 mus musculus
28	186.5	16.3	413	1 ALT3 MOUSE	P22325 cavia porce
29	185	16.2	405	1 ALAS CAVPO	P27359 mus musculus
30	184.5	16.0	418	1 COTR MOUSE	P08185 mus sapien
31	183.5	16.0	405	1 COTR HUMAN	P22323 cavia porce
32	183	16.0	410	1 COTR CAVPO	P17475 rattus norv
33	181.5	15.9	411	1 ALAT RAT	

34	181.5	15.9	417	1 KBP MOUSE	P29621 mus musculus
35	181	15.8	403	1 ALAF CAVPO	P22324 cavia porce
36	181	15.8	416	1 ALAT BOVIN	P34955 bos taurus
37	180	15.7	412	1 ALAT CALCN	O54763 callosciuru
38	178.5	15.6	383	1 CBG RABIT	P23775 cryctolagus
39	178.5	15.6	418	1 CFI6 RAT	P09006 rattus norv
40	175	15.3	380	1 SPB7 HUMAN	O75635 homo sapien
41	175	15.3	409	1 ALAT PAPAN	P01010 papio anubi
42	175	15.3	416	1 ALAT SHEEP	P12725 ovila aries
43	173.5	15.2	436	1 EP45 XENLA	O00387 xenopus lae
44	173	15.1	413	1 ALAF RABIT	P23035 cryctolagus
45	170.5	14.9	418	1 S124 APOSY	Q60396 apodemus ay

## ALIGNMENTS

### RESULT 1

PT16 HUMAN  
ID PT16 HUMAN STANDARD; PRT; 376 AA.  
AC P35237; Q96J44;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 15-FEB-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Placental thrombin inhibitor (Cycloplasmic antiprotease) (CAP)  
DE (Protease inhibitor 6) (PI-6).  
GN SERPINB6 OR P16 OR PTI.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=94022386; PubMed=8415716;  
RA Coughlin P., Sun J., Cerruti L., Salem H.H., Bird P.;  
RT "Cloning and molecular characterization of a human intracellular  
serine proteinase inhibitor";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:9417-9421(1993).  
RN [2]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 67-73 AND 144-149.  
RC TISSUE=Placenta;  
RX MEDLINE=94183847; PubMed=8136380;  
RA Morgenstern K.A., Sprecher C.A., Holth L., Foster D., Grant F.J.,  
Ching A., Kisiel W.;  
RT "Complementary DNA cloning and kinetic characterization of a novel  
intracellular serine proteinase inhibitor: mechanism of action with  
trypsin and factor Xa as model proteinases";  
RL Biochemistry 33:3432-3441(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RA Strausberg R.;  
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Inhibits thrombin.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN SKELETAL MUSCLE. ALSO FOUND  
IN PLACENTA, CARDIAC MUSCLE, LUNG, LIVER, KIDNEY AND PANCREAS.  
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).

CC EMBL; Z22658; CAA80373.1; -  
CC EMBL; S69272; AAB30320.1; -  
CC EMBL; BC001394; AAH01394.1; -  
CC PIR; S35750; S35750.  
CC PIR; A48681; A48681.

KW Serpin; Serine protease inhibitor.  
FT ACT SITE 339 340 REACTIVE BOND (BY SIMILARITY).  
SQ SEQUENCE 374 AA; 42786 MW; E855A033519AC60E CRC64;

Query Match 34.3%; Score 392.5; DB 1; Length 374;  
Best Local Similarity 52.9%; Pred. No. 7.7e-30;  
Matches 83; Conservative 23; Mismatches 34; Indels 17; Gaps 3;

QY 41 KQLIFFPMSISSALAMVFMGAKGNATAQAOMSQLCFSGKIGEDGDIHRGFQSLLVAINRTD 100  
Db 26 RNVFFSPMSISSALAMVFMGAKGNATAQAOMSQLCLYK-----GDIIHRGFQSLLSEVRNG 81  
QY 101 TEYLRLTANGLFGEKSYDFLTGTSCGKFYQATIKQLDFVNDTEKSTRVNSWADTK 160  
Db 82 TQYLRLTANLFGKEKTCDFLPDFKEYCQKFOAELEELSFABETECRKHINDWVAEKTE 141  
QY 161 AWKIQTSLSHLEBPFIASSCYCKACLQPLLVHSI 197  
Db 142 G-----KISEVLDAAGTVDP-----LTKLVLVNAV 165

RESULT 3  
PTI6\_BOVIN STANDARD; PRT; 378 AA.  
ID PTI6\_BOVIN  
AC O02739;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Serine proteinase inhibitor B-43.  
OS Bos taurus (Bovine); Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Eukaryota; Metazoa; Chordata; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Mammalia; Buthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Buthera; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SERPINB8 FROM N.A.  
RC TISSUE=Brain.  
RX MEDLINE=97165893; PubMed=9013786;  
RA Nakaya N., Nishibori M., Kawabata M., Saeiki K.;  
RT "Cloning of a serine proteinase inhibitor from bovine brain:  
RT expression in the brain and characterization of its target  
RT proteinases";  
RL Brain Res. Mol. Brain Res. 42:293-300 (1996).  
CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SERINE PROTEINASES  
CC -!- PRESENT IN THE BRAIN OR EXTRAVASATED FROM THE BLOOD.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- TISSUE SPECIFICITY: BRAIN.  
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; D55670; BAA19875.1; --  
HSSP; P05120; 1BY7.  
InterPro; IPR000215; Serpin.  
Pfam; PF00079; serpin; 1.  
SMART; SM00093; SERPIN; 1.  
PROSITE; PS00284; SERPIN; 1.  
KW Serpin; Serine protease inhibitor.  
FT ACT SITE 343 344 REACTIVE BOND (BY SIMILARITY).  
SQ SEQUENCE 378 AA; 42560 MW; 664F499CCFC263A CRC64;

Query Match 33.9%; Score 388; DB 1; Length 378;  
Best Local Similarity 60.8%; Pred. No. 2.1e-29;  
Matches 73; Conservative 21; Mismatches 26; Indels 0; Gaps 0;

QY 41 KQLIFFPMSISSALAMVFMGAKGNATAQAOMSQLCFSGKIGEDGDIHRGFQSLLVAINRTD 100  
Db 26 RNVFFSPMSISSALAMVFMGAKGNATAQAOMSQLCLYK-----GDIIHRGFQSLLSEVRNG 81  
QY 101 TEYLRLTANGLFGEKSYDFLTGTSCGKFYQATIKQLDFVNDTEKSTRVNSWADTK 160  
Db 82 TQYLRLTANLFGKEKTCDFLPDFKEYCQKFOAELEELSFABETECRKHINDWVAEKTE 141  
QY 161 AWKIQTSLSHLEBPFIASSCYCKACLQPLLVHSI 197  
Db 142 G-----KISEVLDAAGTVDP-----LTKLVLVNAV 165

RESULT 3  
PTI6\_BOVIN STANDARD; PRT; 378 AA.  
ID PTI6\_BOVIN  
AC O02739;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Serine proteinase inhibitor B-43.  
OS Bos taurus (Bovine); Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Eukaryota; Metazoa; Chordata; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Mammalia; Buthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Buthera; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SERPINB8 FROM N.A.  
RC TISSUE=Brain.  
RX MEDLINE=97165893; PubMed=9013786;  
RA Nakaya N., Nishibori M., Kawabata M., Saeiki K.;  
RT "Cloning of a serine proteinase inhibitor from bovine brain:  
RT expression in the brain and characterization of its target  
RT proteinases";  
RL Brain Res. Mol. Brain Res. 42:293-300 (1996).  
CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SERINE PROTEINASES  
CC -!- PRESENT IN THE BRAIN OR EXTRAVASATED FROM THE BLOOD.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- TISSUE SPECIFICITY: BRAIN.  
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; D55670; BAA19875.1; --  
HSSP; P05120; 1BY7.  
InterPro; IPR000215; Serpin.  
Pfam; PF00079; serpin; 1.  
SMART; SM00093; SERPIN; 1.  
PROSITE; PS00284; SERPIN; 1.  
KW Serpin; Serine protease inhibitor.  
FT ACT SITE 343 344 REACTIVE BOND (BY SIMILARITY).  
SQ SEQUENCE 378 AA; 42560 MW; 664F499CCFC263A CRC64;

Query Match 33.9%; Score 388; DB 1; Length 378;  
Best Local Similarity 60.8%; Pred. No. 2.1e-29;  
Matches 73; Conservative 21; Mismatches 26; Indels 0; Gaps 0;

QY 41 KQLIFFPMSISSALAMVFMGAKGNATAQAOMSQLCFSGKIGEDGDIHRGFQSLLVAINRTD 100  
Db 26 RNVFFSPMSISSALAMVFMGAKGNATAQAOMSQLCLYK-----GDIIHRGFQSLLSEVRNG 81  
QY 101 TEYLRLTANGLFGEKSYDFLTGTSCGKFYQATIKQLDFVNDTEKSTRVNSWADTK 160  
Db 82 TQYLRLTANLFGKEKTCDFLPDFKEYCQKFOAELEELSFABETECRKHINDWVAEKTE 141  
QY 161 AWKIQTSLSHLEBPFIASSCYCKACLQPLLVHSI 197  
Db 142 G-----KISEVLDAAGTVDP-----LTKLVLVNAV 165

RESULT 3  
PTI6\_BOVIN STANDARD; PRT; 378 AA.  
ID PTI6\_BOVIN  
AC O02739;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Serine proteinase inhibitor B-43.  
OS Bos taurus (Bovine); Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Eukaryota; Metazoa; Chordata; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Mammalia; Buthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Buthera; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SERPINB8 FROM N.A.  
RC TISSUE=Brain.  
RX MEDLINE=97165893; PubMed=9013786;  
RA Nakaya N., Nishibori M., Kawabata M., Saeiki K.;  
RT "Cloning of a serine proteinase inhibitor from bovine brain:  
RT expression in the brain and characterization of its target  
RT proteinases";  
RL Brain Res. Mol. Brain Res. 42:293-300 (1996).  
CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SERINE PROTEINASES  
CC -!- PRESENT IN THE BRAIN OR EXTRAVASATED FROM THE BLOOD.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- TISSUE SPECIFICITY: BRAIN.  
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; D55670; BAA19875.1; --  
HSSP; P05120; 1BY7.  
InterPro; IPR000215; Serpin.  
Pfam; PF00079; serpin; 1.  
SMART; SM00093; SERPIN; 1.  
PROSITE; PS00284; SERPIN; 1.  
KW Serpin; Serine protease inhibitor.  
FT ACT SITE 343 344 REACTIVE BOND (BY SIMILARITY).  
SQ SEQUENCE 378 AA; 42560 MW; 664F499CCFC263A CRC64;

Query Match 33.9%; Score 388; DB 1; Length 378;  
Best Local Similarity 60.8%; Pred. No. 2.1e-29;  
Matches 73; Conservative 21; Mismatches 26; Indels 0; Gaps 0;

QY 41 KQLIFFPMSISSALAMVFMGAKGNATAQAOMSQLCFSGKIGEDGDIHRGFQSLLVAINRTD 100  
Db 26 RNVFFSPMSISSALAMVFMGAKGNATAQAOMSQLCLYK-----GDIIHRGFQSLLSEVRNG 81  
QY 101 TEYLRLTANGLFGEKSYDFLTGTSCGKFYQATIKQLDFVNDTEKSTRVNSWADTK 160  
Db 82 TQYLRLTANLFGKEKTCDFLPDFKEYCQKFOAELEELSFABETECRKHINDWVAEKTE 141  
QY 161 AWKIQTSLSHLEBPFIASSCYCKACLQPLLVHSI 197  
Db 142 G-----KISEVLDAAGTVDP-----LTKLVLVNAV 165

RESULT 3  
PTI6\_BOVIN STANDARD; PRT; 378 AA.  
ID PTI6\_BOVIN  
AC O02739;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Serine proteinase inhibitor B-43.  
OS Bos taurus (Bovine); Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Eukaryota; Metazoa; Chordata; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Mammalia; Buthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Buthera; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SERPINB8 FROM N.A.  
RC TISSUE=Brain.  
RX MEDLINE=97165893; PubMed=9013786;  
RA Nakaya N., Nishibori M., Kawabata M., Saeiki K.;  
RT "Cloning of a serine proteinase inhibitor from bovine brain:  
RT expression in the brain and characterization of its target  
RT proteinases";  
RL Brain Res. Mol. Brain Res. 42:293-300 (1996).  
CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SERINE PROTEINASES  
CC -!- PRESENT IN THE BRAIN OR EXTRAVAS

```

Db 25 KVVFTSPLSISALAMVLGAKGNTAAQCTLSLNKSGGGEDVHQGFQNLSEVNRD 84
QY 101 TEYLVRTANGLEKSYDPLTGFTDCGKFKYQATIKOLDVNDTEKSTTRVNSWADTK 160
Db 85 TOYLLRTANRLGKTYDPLSFKSCHFKFYQAEEMELDFVSAEQSRKHINTWVAEKTE 144

RESULT 4
PT16_MOUSE
ID PT16_MOUSE STANDARD; PRT; 378 AA.
AC Q60854;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Placental thrombin inhibitor (protease inhibitor 6) (PI-6);
GN SERPINB6 OR SPI3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=95332310; PubMed=7608171;
RA Sun J., Rose J.B., Bird P.;
RT "Gene structure, chromosomal localization, and expression of the murine homologue of human proteinase inhibitor 6 (PI-6) suggests divergence of PI-6 from the ovalbumin serpins.";
RL J. Biol. Chem. 270:16089-16096(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Inhibits thrombin.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
DR EMBL; U25844; AAA79684.1; -
DR EMBL; BC006766; AAH06766.1; -
DR HSSP; P05120; 1BY7.
DR MGD; MGI:103123; Serpinb6.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT SITE 343 344 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 378 AA; 42599 MW; 4B0F5E1A030BBD6 CRC64;

Query Match 32.4%; Score 370.5; DB 1; Length 378;
Best Local Similarity 47.2%; Pred. No. 9.6e-28;
Matches 77; Conservative 24; Mismatches 43; Indels 19; Gaps 3;

QY 31 PFEKARG-----KQLKLIFFPMSTISSALAMVFMGAKGNTAAQMSQALCFSKI 78
Db 3 PLQEAANGTFALNLLKILGEDSSKNVFLSPMSISSALAMVFMGAKGNTAAQALADKC 62
QY 79 GGE-DGDIHRGQSLLVAINRTDTBYVLRNTANGLFGKSYDPLTGFTDCGKFKYQATIKQ 137
Db 63 SONGGGVHQGFQSSLLTEVNTKGTQYLLRTANRLFGDKTCLLASFKDCLKFEAELEE 122
QY 138 LDFVNDTEKSTTRVNSWADTKKAWKIQTSLSHLEEPGIASS 180
Db 123 LDFQATGESRQHINTWAKTI-----EDKIKEVLSPTGVNS 159

```

```

RESULT 5
SPB9_HUMAN
ID SPB9_HUMAN STANDARD; PRT; 376 AA.
AC P50453;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE Cytoplasmic antiprotease 3 (CAP3) (CAP-3) (Protease inhibitor 9) (Serpin B9).
GN SERPINB9 OR PI9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=96102039; PubMed=8530382;
RA Sprecher C.A., Morgenstern K.A., Mathewes S., Dahlen J.R.,
RA Schrader S.K., Foster D.C., Kisiel W.;
RT "Molecular cloning, expression, and partial characterization of two novel members of the ovalbumin family of serine proteinase inhibitors.";
RL J. Biol. Chem. 270:29854-29861(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta, and Bone marrow;
RX MEDLINE=97066975; PubMed=8910377;
RA Sun J., Bird C.H., Sutton V., McDonald L., Coughlin P.B., Jong T.A.,
RA Trapani J.A., Bird P.I.;
RT "A cytosolic granzyme B inhibitor related to the viral apoptotic regulator cytokine response modifier A is present in cytotoxic lymphocytes.";
RL J. Biol. Chem. 271:27802-27809(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: GRANZYME B INHIBITOR.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
DR EMBL; L40378; AAC41940.1; -
DR EMBL; U71364; AAC50793.1; -
DR EMBL; BC002538; AAH02538.1; -
DR HSSP; P05120; 1BY7.
DR Genew; HGNC:8955; SERPINB9.
DR MIM; 601799; -
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT SITE 340 341 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 376 AA; 42403 MW; ECB8A0869B939753 CRC64;

Query Match 27.9%; Score 320; DB 1; Length 376;
Best Local Similarity 46.3%; Pred. No. 6e-23;
Matches 68; Conservative 28; Mismatches 45; Indels 6; Gaps 2;

QY 47 PMSISSALAMVFMGAKGNTAAQMSQALCFSKIGGDDIHRGQSLLVAINRTDTYVLR 106

```

Wed May 28 14:44:10 2003

Db 32 PVSISALAMVLLGAKNGTATQMAQALSIN-----TEEDIHRAFOQLLTVNKGCTQYLLR 87  
 QY 107 TANGLFGEKSDYFLGFTDCCKFYQATIKOLDVFNDEKSTTRVNSWADKT--KAWKI 164  
 Db 88 YANRLFGEKTCQFLSTFKESCLQFYHABLKSLFIRAABESRKHINTWVSKKTEGKEEL 147  
 QY 165 IOTSLSHLEPGIASSCYCKACLSP 191  
 Db 148 LPGSGIDAEIRLVNLVNAIFYFGKWNKP 174

RESULT 6  
 SCC2\_HUMAN STANDARD; PRT; 390 AA.  
 AC P48594;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Squamous cell carcinoma antigen 2 (SCCA-2) (Leupin).  
 GN SERPINB4 OR SCCA2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95241462; PubMed=7724531;  
 RA Schneider S.S., Schick C., Fish K.E., Miller E., Pena J.C.,  
 RA Treter S.D., Hui S.M., Silverman G.A.;  
 RT "A serine proteinase inhibitor locus at 18q21.3 contains a tandem  
 duplication of the human squamous cell carcinoma antigen gene."  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3147-3151(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96013887; PubMed=7589435;  
 RA Barnes R.C., Worrall D.M.;  
 RA "Identification of a novel human serpin gene; cloning sequencing and  
 RT expression of leupin."  
 RL FEBS Lett. 373:61-65(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21167379; PubMed=11267667;  
 RA Hamada K., Shinomiya H., Asano Y., Kihana T., Iwamoto M., Hanakawa Y.,  
 RA Hashimoto K., Hirose S., Ito M.;  
 RT "Molecular cloning of human squamous cell carcinoma antigen 1 gene and  
 RT characterization of its promoter."  
 RL Biochim. Biophys. Acta 1518:124-131(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Bone marrow;  
 RA Strausberg R.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MAY ACT AS A PROTEASE INHIBITOR TO MODULATE THE HOST  
 CC -1- IMMUNE RESPONSE AGAINST TUMOR CELLS.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- TISSUE SPECIFICITY: SQUAMOUS CELLS.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U19576; AAA92602.1; JOINED.  
 DR EMBL; U19570; AAA92602.1; JOINED.  
 DR EMBL; U19571; AAA92602.1; JOINED.  
 DR EMBL; U19572; AAA92602.1; JOINED.  
 DR EMBL; U19574; AAA92602.1; JOINED.  
 DR EMBL; U19575; AAA92602.1; JOINED.  
 DR EMBL; U19557; AAA97553.1; JOINED.

DR EMBL; X89015; CAA61420.1; --  
 DR EMBL; AB035089; BAB21525.1; --  
 DR EMBL; BC017401; AAH17401.1; --  
 DR HSP; P01008; IATH.  
 DR Genew; HGNC:10570; SERPINB4.  
 DR MIM; 600518; --  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Serpin; Serine protease inhibitor.  
 FT ACT\_SITE 354 355 REACTIVE BOND.  
 SQ SEQUENCE 390 AA; 44854 MW; 04E213CD892587D5 CRC64;

Query Match 23.8%; Score 272; DB 1; Length 390;  
 Best Local Similarity 38.6%; Pred. No. 2.3e-18;  
 Matches 54; Conservative 27; Mismatches 47; Indels 12; Gaps 1;  
 QY 32 FEKARGKQLQLFFPMSSISSALAMVFMGAKNTAAQMSQALCFSGKIG----- 79  
 Db 16 FQPRKSKENNIFYSPISITSLGWLGLGAKNTAQISKVLHFDQVTENTTEKAATYHV 75  
 QY 80 GEDGDHRRGFSLLVAINRTDTYVLRTANGLFGKSYDELFTGTCGKFGYQATIKOLD 139  
 Db 76 DRSGNVHHQFQKLLTFNFKSTDAVELKIANKLFGEKTYQFLQBYLDAIKKPYQTSVESTD 135  
 QY 140 FVNDTEKSTTRVNSWADKT 159  
 Db 136 FANAPESRKRKINSWVESQT 155

RESULT 7  
 SCC1\_HUMAN STANDARD; PRT; 390 AA.  
 AC P29508; Q96J21;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Squamous cell carcinoma antigen 1 (SCCA-1) (Protein T4-A).  
 GN SERPINB3 OR SCCA1 OR SCCA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=92068241; PubMed=1958219;  
 RA Suminami Y., Kishi F., Sekiguchi K., Kato H.;  
 RT "Squamous cell carcinoma antigen is a new member of the serine  
 RT protease inhibitors."  
 RL Biochem. Biophys. Res. Commun. 181:51-58(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95241462; PubMed=7724531;  
 RA Schneider S.S., Schick C., Fish K.E., Miller E., Pena J.C.,  
 RA Treter S.D., Hui S.M., Silverman G.A.;  
 RT "A serine proteinase inhibitor locus at 18q21.3 contains a tandem  
 RT duplication of the human squamous cell carcinoma antigen gene."  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3147-3151(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Lung;  
 RA Strausberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MAY ACT AS A PROTEASE INHIBITOR TO MODULATE THE HOST  
 CC -1- IMMUNE RESPONSE AGAINST TUMOR CELLS.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- TISSUE SPECIFICITY: SQUAMOUS CELLS.  
 CC -1- DEVELOPMENTAL STAGE: ITS EXPRESSION IS CLOSELY RELATED TO CELLULAR  
 CC DIFFERENTIATION IN BOTH NORMAL AND MALIGNANT SQUAMOUS CELLS.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; S6696; AAB20405.1; -  
 DR EMBL; U19556; AAA97552.1; -  
 DR EMBL; U19558; AAA86317.1; -  
 DR EMBL; U19559; AAA86317.1; JOINED.  
 DR EMBL; U19560; AAA86317.1; JOINED.  
 DR EMBL; U19562; AAA86317.1; JOINED.  
 DR EMBL; U19565; AAA86317.1; JOINED.  
 DR EMBL; U19567; AAA86317.1; JOINED.  
 DR EMBL; U19562; AAA86316.1; -  
 DR EMBL; U19559; AAA86316.1; JOINED.  
 DR EMBL; U19560; AAA86316.1; JOINED.  
 DR EMBL; BC005224; AAA805224.1; -  
 DR PIR; JT0966; JT0966.  
 DR HSSP; P01008; IATH.  
 DR Genew; HGNC:10569; SERPINB3.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 DR Serpin; Serine protease inhibitor.  
 DR ACT\_SITE 354 355 REACTIVE BOND.  
 FT CONFLICT 357 357 T -> A (IN REF. 1).  
 SQ SEQUENCE 390 AA; 44564 MW; E5F27F986C752CFA CRC64;

Query Match 23.28; Score 266; DB 1; Length 390;  
 Best Local Similarity 37.94; Pred. No. 8.4e-18;  
 Matches 53; Conservative 28; Mismatches 47; Indels 12; Gaps 1;  
 32 FEKARGKQLKQIFPPMSISSALAVPMGAKNTAAQMSQALCFSGKIG----- 79  
 16 FQPRKSKENNIFYSPTSISALGVLLGAKNTAQIKVLFHQVNTTNGKAATYHV 75  
 80 GEGDIIHRGQSLLVAINRTDTEYLVRTANGLFGEKSYDFLTGTFDSCGFYQATIKQLD 139  
 76 DRGNVHHQFKLLTEFNKSTDAYELKIANKLFGEKTYLFLQBYLDAIKKFQTSVESVD 135  
 140 FVNDTEKSTTRVNSWADKT 159  
 136 FANAPESRKINSVESQT 155.

RESULT 8  
 ID ILEU\_HORSE STANDARD; PRT; 379 AA.  
 AC P05619;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Leukocyte elastase inhibitor (LEI).  
 GN SERPINB1 OR ELANH2.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93319507; PubMed=7687128;  
 RA Kordula T., Dubin A., Schooltink H., Koj A., Heinrich P.C.,  
 RA Rose-John S.;  
 RT "Molecular cloning and expression of an intracellular serpin: an  
 RT elastase inhibitor from horse leukocytes.";  
 RL Biochem. J. 293:187-193 (1993).  
 RN [2]  
 RP SEQUENCE.  
 RC TISSUE=Leukocyte;

RA MEDLINE=92202200; PubMed=1551869;  
 RA Dubin A., Travis J., Enghild J.J., Potempa J.;  
 RT "Equine leukocyte elastase inhibitor. Primary structure and  
 RT identification as a chymosin-binding protein.";  
 RL J. Biol. Chem. 267:6576-6583 (1992).  
 RN [3]  
 RP SEQUENCE OF 343-362.  
 RX MEDLINE=8213423; PubMed=3366785;  
 RX Potempa J., Dubin A., Warerek W., Travis J.;  
 RT "An elastase inhibitor from equine leukocyte cytosol belongs to the  
 RT serpin superfamily. Further characterization and amino acid sequence  
 RT of the reactive center.";  
 RL J. Biol. Chem. 263:7364-7369 (1988).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).  
 RX MEDLINE=92389327; PubMed=1518052;  
 RX Baumann U., Bode W., Huber R., Travis J., Potempa J.;  
 RT "Crystal structure of cleaved equine leukocyte elastase inhibitor  
 RT determined at 1.95-A resolution.";  
 RL J. Mol. Biol. 226:1207-1218 (1992).  
 CC -!- FUNCTION: THIS INHIBITOR IS THOUGHT TO BE INVOLVED IN THE CONTROL  
 CC OF INTRACELLULAR PROTEIN TURNOVER. BINDS TO THYMOSIN BETA-4.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M91161; AAA97513.1; -  
 DR PIR; A28060; A28060.  
 DR PIR; A37276; A37276.  
 DR PIR; A42421; A42421.  
 DR PIR; S34062; S34062.  
 DR PDB; 1HLE; 31-JAN-94.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 DR Serpin; Serine protease inhibitor; 3D-structure.  
 FT MOD\_RES 1 1 BLOCKED (WITH AN UNKNOWN CHEMICAL GROUP).  
 FT ACT\_SITE 344 345 REACTIVE BOND.  
 FT HELIX 2 22  
 FT STRAND 28 30  
 FT HELIX 32 44  
 FT TURN 45 45  
 FT HELIX 48 58  
 FT TURN 59 59  
 FT HELIX 60 62  
 FT TURN 64 65  
 FT HELIX 66 77  
 FT TURN 78 78  
 FT STRAND 85 95  
 FT TURN 96 97  
 FT HELIX 102 112  
 FT STRAND 115 119  
 FT TURN 121 123  
 FT HELIX 125 139  
 FT TURN 140 142  
 FT TURN 150 151  
 FT TURN 155 156  
 FT STRAND 157 171  
 FT HELIX 177 179  
 FT STRAND 181 186  
 FT STRAND 192 209  
 FT HELIX 210 212  
 FT TURN 213 213  
 FT STRAND 214 221  
 FT TURN 222 222

QY 47 PMSISSALAMVFMGAKNTAAQMSQALCFKIGGEDGDIHRGFSQSLVAINRTDTEYVLR 106  
 Db 32 PFSISSALAMILLTRGNTAQAQSKALHFDTV---KDIHSRFSQSLNADINKCGASYILK 87  
 QY 107 TANGLFGEKSYDFLTGTDSCGKGYQATIKQLDFVNDTEKSTTRVNSVWADTKAWKIIO 166  
 Db 88 LANRLFGKTYHFLPEFLASTOKTYGAELASVDFLRASEEARKAINWVKEQTEG----- 142  
 QY 167 TSLSHLEEPGIASSCYCKACLSQPLLVHSI 197  
 Db 143 -KIPELLASGWDSA-----TKLVLVNAI 165

RESULT 10  
 ILEU HUMAN STANDARD; PRT; 379 AA.  
 ID ILEU HUMAN AC P30740;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Leukocyte elastase inhibitor (LEI) (Monocyte/neutrophil elastase  
 DE inhibitor) (M/NEI) (EI).  
 GN SERPINB1 OR ELANH2 OR P12.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 QY 1 SEQUENCE FROM N.A.  
 RP MEDLINE=92302296; PubMed=1376927;  
 RX Remold-O'Donnell E., Chin J., Alberts M.;  
 RA "Sequence and molecular characterization of human monocyte/neutrophil  
 RT elastase inhibitor";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:5635-5639(1992).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92302296; PubMed=1376927;  
 RA Remold-O'Donnell E., Chin J., Alberts M.;  
 RT "Structure and sequence of human M/NEI (monocyte/neutrophil elastase  
 RT inhibitor), an Ov-serpin family gene";  
 RL Gene 213:179-187(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Cervix;  
 RA Strausberg R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 178-185; 204-210 AND 364-371.  
 RX TISSUE=Keratinocytes;  
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,  
 RT Vandekerckhove J.;  
 RL "Microsequences of 145 proteins recorded in the two-dimensional gel  
 RL Electrophoresis 13:960-969(1992).  
 CC -!- FUNCTION: REGULATES THE ACTIVITY OF THE NEUTROPHIL PROTEASES  
 CC ELASTASE, CATHEPSIN G AND PROTEINASE-3.  
 CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; M91056; -; NOT ANNOTATED\_CDS.  
 CC EMBL; AF053630; AAC31394.1; -;  
 CC EMBL; BC009015; A0409015.1; -;  
 CC PIR; S27383; S27383.  
 CC DR  
 CC DR  
 CC DR

FT STRAND 226 233  
 FT HELIX 244 248  
 FT TURN 249 249  
 FT TURN 259 259  
 FT HELIX 261 263  
 FT STRAND 265 274  
 FT STRAND 276 283  
 FT STRAND 285 291  
 FT TURN 292 292  
 FT TURN 295 297  
 FT TURN 299 301  
 FT TURN 301 308  
 FT STRAND 313 326  
 FT STRAND 330 343  
 FT STRAND 350 353  
 FT STRAND 358 364  
 FT TURN 365 368  
 FT STRAND 369 376  
 SQ SEQUENCE 379 AA; 43006 MW; 72DED1999371427A CRC64;  
 Query Match 22.4%; Score 257; DB 1; Length 379;  
 Local Similarity 46.5%; Pred. No. 5.8e-17; Mismatches 38; Indels 4; Gaps 1;  
 Matches 53; Conservative 19;

QY 47 PMSISSALAMVFMGAKNTAAQMSQALCFKIGGEDGDIHRGFSQSLVAINRTDTEYVLR 106  
 Db 32 PLSISSALAMILLTRGNTAQAQSKALHFDTV---EDIHSRFSQSLNADINKCGASYILK 87  
 QY 107 TANGLFGEKSYDFLTGTDSCGKGYQATIKQLDFVNDTEKSTTRVNSVWADTK 160  
 Db 88 LANRLFGKTYHFLPEFLASTOKTYGAELASVDFOQAPEDARKKEINWVKGQTE 141

RESULT 9  
 ILEU PIG STANDARD; PRT; 378 AA.  
 ID ILEU PIG AC P80229;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Leukocyte elastase inhibitor (LEI) (Leukocyte neutral proteinase  
 DE inhibitor) (LNPI).  
 GN SERPINB1 OR ELANH2.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.  
 OC NCBI\_TaxID=9823;  
 QY 1 SEQUENCE.  
 RP MEDLINE=94039085; PubMed=7901009;  
 RA Teschauer W.F., Mentele R., Sommerhoff C.P.;  
 RT "Primary structure of a porcine leukocyte serpin";  
 RL Eur. J. Biochem. 217:519-526(1993).  
 CC -!- FUNCTION: REGULATES THE ACTIVITY OF THE NEUTROPHIL PROTEASE  
 CC ELASTASE, CATHEPSIN G AND PROTEINASE-3 (By similarity).  
 CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC HSP; P05619; 1HLE.  
 DR InterPro: IPR000215; Serpin.  
 DR Pfam: PF00079; serpin; 1.  
 DR SMART: SM00093; SERPIN; 1.  
 DR PROSITE: PS00284; SERPIN; 1.  
 KW Serpin; Serine protease inhibitor.  
 FT MOD RES 1 1 BLOCKED.  
 FT DISULFID 80 80 INTERCHAIN (PROBABLE).  
 FT ACT SITE 343 344 REACTIVE BOND (BY SIMILARITY).  
 SQ SEQUENCE 378 AA; 42512 MW; 3CA7C65F11E508CA CRC64;  
 Query Match 22.3%; Score 255.5; DB 1; Length 378;  
 Best Local Similarity 39.1%; Pred. No. 8e-17; Mismatches 48; Indels 17; Gaps 3;  
 Matches 59; Conservative 27;

DR HSP: P05619; 1HLU.  
DR Aarhu/Ghent-2DPAGE; 4314; IBF.  
DR Genew; HGNC:3311; SERPINB1.  
DR MIM; 130135; -.  
DR InterPro; IPR00215; Serpin.  
DR Pfam; PF00079; serpin.1.  
DR SMART; SM00093; SERPIN; 1.  
DR PROSITE; PS00284; SERPIN; 1.  
KW Serpin; Serine protease inhibitor.  
FT ACT SITE 344 345 REACTIVE BOND (BY SIMILARITY).  
SQ SEQUENCE 379 AA; 42741 MW; BAAE08DFC8DCD3 CRC64;

Query Match  
Best Local Similarity 21.0%; Score 240; DB 1; Length 379;  
Matches 49; Conservative 23; Mismatches 38; Indels 4; Gaps

QY 47 PMSISSALAMVFMGKGNVTAQMQLCFSGKIGGEDGDIHRFGQSLLVAINRTDTYVLR 106  
D 32 PFSISSAMVFLGTRGNTAAQLSKTFHTV-----EVHSRFSQSLNADINKKGASYILK 87  
QY 107 TANGFLYGEKSKYDGLTGFTDSCGKYQYATIKOLDPVNDTEKSTRVNSWVADKTK 160  
D 88 LANRLYGEKTYNLFUFLVSTQKTYGADLASVDFQHASDARKTINQWVKGQTE 141

RESULT 11

SD13 HUMAN  
ID SB13 HUMAN STANDARD; PRT; 391 AA.  
AC Q9UIV8; Q9UBW1; Q9UKG0; Q9HCX1;  
DT 16-OCT-2001 (Rel. 40, Created)  
DI 16-OCT-2001 (Rel. 40, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hurpin (HaCat UV-repressible serpin) (Protease inhibitor 13)  
DE (Headpin) (Serpin B13).  
GN SERPINB13 OR P113.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A. (ISOFORM 1).  
TISSUE=Skin;  
RC MEDLINE=99443898; PubMed=10512713;  
RX Abts H.F., Welles T., Mirmohammadsadeh A., Koehler K., Michel G.,  
RA Ruzicka T.;  
RT "Cloning and characterization of hurpin (Protease Inhibitor 13): a new  
RT skin specific, UV-repressible serine proteinase inhibitor of the  
RL ovalbumin serpin family";  
RL J. Mol. Biol. 293:29-39(1999).  
[2]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=99458661; PubMed=10527881;  
RT Spring P., Nakashima T., Frederick M., Henderson V., Clayman G.;  
RT "Identification and cDNA cloning of headpin, a novel differentially  
RT expressed serpin that maps to chromosome 18q.";  
RL Biochem. Biophys. Res. Commun. 264:299-304(1999).  
[3]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=20461850; PubMed=11004515;  
RA Nakashima T., Pak S.C., Silverman G.A., Spring P.M., Frederick M.J.,  
RA Clayman G.L.;  
RT "Genomic cloning, mapping, structure and promoter analysis of HEADPIN,  
RT a serpin which is down-regulated in head and neck cancer cells";  
RL Biochim. Biophys. Acta 1492:441-446(2000)..  
[4]  
RN SEQUENCE FROM N.A. (ISOFORM 2).  
RP TISSUE=Placenta;  
RX MEDLINE=21213570; PubMed=11313015;  
RA Abts H.F., Welles T., Scheuring S., Scott F.L., Irving J.A., Michel G.,  
RA Bird P.I., Ruzicka T.;  
RT "Sequence, organization, chromosomal localization and alternative  
RT splicing of the human serine proteinase inhibitor gene hurpin (P113),  
RT which is up-regulated in psoriasis";

123 CEMSCCKEYCATIKOI.DEVNDTEKSTTRVNSWVADKTK 160



Db 140 EYRLCQYSSSEPOAVDFLECAEARKKINSWVKTK 178

RESULT 13

PA12\_MOUSE

ID PA12\_MOUSE STANDARD; PRT; 415 AA.

AC P12388; Q35687; Q9QWZ5; Q9QWQ0; Q9QWP9; Q9QWP8; Q9QWP7;

AC Q9QWP6;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Plasminogen activator inhibitor-2, macrophage (PAI-2).

GN SERPINB2 OR PAI2 OR PLANH2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

ON NCBI\_TaxID=10090;

RN [1]

RX STRAIN=AKR; TISSUE=Myeloid;

RX MEDLINE=90059920; PubMed=2583099;

RA Belin D., Wohlwend A., Schleuning W.-D., Kruithof E.K.O.,

RA Vassalli J.-D.,

RT "Facultative polypeptide translocation allows a single mRNA to encode

RT the secreted and cytosolic forms of plasminogen activators inhibitor

RT 2.";

RL EMBO J. 8:3287-3294 (1989).

RN [2]

RN SEQUENCE FROM N.A.

RC STRAIN=AKR; TISSUE=Myeloid;

RC Submitted (SEP-1989) to the EMBL/GenBank/DBJ databases.

RN [3]

RN SEQUENCE FROM N.A., AND VARIANTS GLU-239; TYR-258; ILE-259, AND

RP ARG-284.

RC STRAIN=129; TISSUE=Liver;

RA Belin D., Tapparel C., Sappino N., Silva P.;

RT "Sequence and tissue-specific expression of the murine PAI-2 gene.";

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE OF 1-34; 54-74; 89-115; 124-156; 163-237 AND 265-302 FROM

RP N.A.

RC STRAIN=129/Sv;

RC MEDLINE=99110951; PubMed=9892694;

RA Dougherty K.M., Pearson J.M., Yang A.Y., Westrick R.J., Baker M.S.,

RA Ginsburg D.;

RT "The plasminogen activator inhibitor-2 gene is not required for normal

RL murine development or survival.";

RL Proc. Natl. Acad. Sci. U.S.A. 96:686-691 (1999).

CC -!- FUNCTION: PAI-2 inhibits urokinase-type plasminogen activator. The

CC monocyte derived PAI-2 is distinct from the endothelial cell-

CC derived PAI-1. Not required for normal murine development or

CC survival.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC OR EXTRACELLULAR.

CC -!- PTM: THE SIGNAL SEQUENCE IS NOT CLEAVED.

CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X16490; CAA34507.1; -

DR EMBL; AJ000386; CAA04047.1; -

DR EMBL; AJ000384; CAA04047.1; JOINED.

DR EMBL; AJ000385; CAA04047.1; JOINED.

DR EMBL; AF069683; AAD08915.1; -

DR EMBL; AF069685; AAD08916.1; -

DR EMBL; AF069684; AAD08916.1; JOINED.

DR EMBL; AF069687; AAD08917.1; -

DR EMBL; AF069686; AAD08917.1; JOINED.

DR EMBL; AF069689; AAD08918.1; -

DR EMBL; AF069688; AAD08918.1; JOINED.

DR EMBL; AF069690; AAD08919.1; -

DR EMBL; AF069692; AAD08920.1; -

DR EMBL; AF069691; AAD08920.1; JOINED.

DR EMBL; AF069694; AAD08921.1; -

DR EMBL; AF069693; AAD08921.1; JOINED.

DR PIR; S09616; S09616.

DR PIR; S20047; S20047.

DR HSSP; P05120; 1BY7.

DR MGD; MGI:97609; Serpinb2.

DR InterPro; IPR000215; Serpin.

DR Pfam; PF00079; serpin; 1.

DR SMART; SM00093; SERPIN; 1.

DR PROSITE; PS00284; SERPIN; 1.

KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;

KW Glycoprotein; Signal; Polymorphism.

FT SIGNAL 1 ? NOT CLEAVED.

FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT ACT SITE 380 381 REACTIVE BOND.

FT VARIANT 239 239 A -> E.

FT VARIANT 258 258 H -> Y.

FT VARIANT 259 259 T -> I.

FT VARIANT 284 284 S -> R.

FT CONFLICT 55 55 A -> T (IN REF. 3 AND 4).

SQ SEQUENCE 415 AA; 46291 MW; C4FC78C5C3CA7D8B CRC64;

Query Match 18.9%; Score 216.5; DB 1; Length 415;

Best Local Similarity 32.7%; Pred. No. 4.5e-13;

Matches 54; Conservative 23; Mismatches 53; Indels 35; Gaps 3;

Qy 29 IKPEKARGKQLQKLIFFPMSSISSALAMVFMGAKGNTAAQMSQALCFSGK----- 79

Db 16 LKQIEKSNSTQ--NIFSPWSISSTLAIVLLGAGGNTQQMAKVLQFNEIGSVGITRNP 73

Qy 80 -----GEDGD-IHRGFOSLLVAINRTDTEVLRVLTANGLFGEK 115

Db 74 ENFSGCDFAQIQKENYPSAILQAQAGDKIHSAFSSLSSTINTPOGDYLLSANKLFGEK 133

Qy 116 SYDFLGTGFTDSCGKGYQATIKQLDFVNTDEKSTTRVNGSVADTK 160

Db 134 SARFKEEYQLSKKYSTPEPEAVDFLECAEAREKINSWVKTK 178

RESULT 14

SB10\_HUMAN

ID SB10\_HUMAN STANDARD; PRT; 397 AA.

AC P48595;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Bompain (Protease inhibitor 10) (Serpin B10).

GN SERPINB10 OR PI10.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RX SEQUENCE FROM N.A.

RC TISSUE=Bone marrow.

RC MEDLINE=96070759; PubMed=7592909;

RA Riewald M., Schlegel R.R.,

RT "Molecular cloning of bompain (protease inhibitor 10), a novel human

RT serpin that is expressed specifically in the bone marrow.";

EL J. Biol. Chem. 270:26754-26757 (1995).

CC -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF PROTEASE

CC ACTIVITIES DURING HEMATOPOIESIS.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

```
--!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE BONE MARROW.
--!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; U35459; AAC50282.1; -.
DR HSSP; P05619; IHLE; SERPINB10.
DR Genew; HGNC:8942; SERPINB10.
DR MIM; 602058; -.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT_SITE 362 363 REACTIVE BOND (BY SIMILARITY).
SV SEQUENCE 397 AA; 45402 MW; 8CE01246867154DF CRC64;
-----
Query Match          18.6%; Score 213.5; DB 1; Length 397;
Best Local Similarity 32.2%; Pred. No. 8.3e-13;
Matches 47; Conservative 34; Mismatches 44; Indels 21; Gaps 2;
-----
QY 33 EKARGKQLKLIFFPMISALLAWFMGAKNGNTAAQMSQLCFSKIGG----- 80
DB 21 ESAQG---KNIFFSSWISLSLTIVILGAKGTITNAQMAQVLQFNRDQGVCKDPESEKKR 77
QY 81 -----EDGDTHRFQSILVAIRNTRDETVYLRTANGLFGEKSYDFLTGFTDCSGKFVAT 134
DB 78 MEFLNSSEEIHSDFQTLSEILLKPNDLDLLKTANAIFYEKTYPHNKYLEDKMTYFGAE 137
QY 135 IKQLDFVNDETEKSTRVNSWADTK 160
DB 138 PQPVNFVEASQIRKDINSWVERQTE 163
-----
RESULT 15
PA12 RAT PAI2 RAT STANDARD; PRT; 416 AA.
AC P29524;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen activator inhibitor-2, type A (PA12A).
GN SERPINB2 OR PA12 OR PLANH2.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Grundmann U., Rein T.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PAI-2 INHIBITS UKONASE-TYPE PLASMINOGEN ACTIVATOR.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC OR EXTRACELLULAR.
CC (BY SIMILARITY).
CC -!- PTM: THE SIGNAL SEQUENCE IS NOT CLEAVED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; X64563; CAA45864.1; -.
DR PIR; S19896; S19896.
```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 10:19:47 ; Search time 84 Seconds  
(without alignments)

527.382 Million cell updates/sec

Title: US-09-912-628-7

Perfect score: 1145

Sequence: 1 HELRSWAARRTGAHRGCS.....SIPKNSPVTPHGMWXPSSL 215

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Spaced: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	377.5	33.0	377	11 O08804	O08804 mus musculus
2	347.5	30.3	385	11 Q8VHQ0	Q8VHQ0 mus musculus
3	344	30.0	374	11 O08800	O08800 mus musculus
4	272	23.8	369	4 Q9BYF7	Q9BYF7 homo sapien
5	266	23.2	338	4 Q9BYF8	Q9BYF8 homo sapien
6	259	22.6	374	11 O08757	O08757 mus musculus
7	251.5	22.0	382	11 Q8VRP7	Q8VRP7 mus musculus
8	249.5	21.8	377	11 O08806	O08806 mus musculus
9	248.5	21.7	377	11 Q9DAZ7	Q9DAZ7 mus musculus
10	248.5	21.7	377	11 Q8VHQ1	Q8VHQ1 mus musculus
11	245	21.4	387	11 Q9DIQ5	Q9DIQ5 mus musculus
12	243	21.2	379	11 Q9DOS8	Q9DOS8 mus musculus
13	243	21.2	379	11 Q9D154	Q9D154 mus musculus
14	240	21.0	377	11 Q9DAV6	Q9DAV6 mus musculus
15	237	20.7	379	11 Q9D7S8	Q9D7S8 mus musculus
16	225.5	19.7	388	11 Q922G2	Q922G2 mus musculus

17	210	18.3	410	13	073790	073790 gallus gall
18	204	17.8	359	11	Q9D6A7	Q9D6A7 mus musculus
19	196.5	17.2	386	11	Q9D1E7	Q9D1E7 mus musculus
20	194	16.9	421	6	O46519	O46519 equus caball
21	191.5	16.7	413	11	Q8VC20	Q8VC20 mus musculus
22	190.5	16.6	410	11	Q8VH5	Q8VH5 mus musculus
23	190.5	16.6	413	11	Q8VC41	Q8VC41 mus musculus
24	190.5	16.6	413	11	Q91V74	Q91V74 mus musculus
25	188.5	16.5	418	11	Q03734	Q03734 mus musculus
26	187.5	16.4	388	11	Q9CQV3	Q9CQV3 mus musculus
27	187.5	16.4	388	11	Q91Z12	Q91Z12 mus musculus
28	187.5	16.4	418	11	Q62257	Q62257 mus musculus
29	187	16.3	423	11	Q9D7P9	Q9D7P9 mus musculus
30	186.5	16.3	380	11	Q920J5	Q920J5 rattus norv
31	186.5	16.3	413	11	Q91XB8	Q91XB8 mus musculus
32	186.5	16.3	418	11	Q62258	Q62258 mus musculus
33	186.5	16.3	418	11	Q91WP6	Q91WP6 mus musculus
34	186.5	16.3	425	11	Q91XC1	Q91XC1 mus musculus
35	185	16.2	406	11	Q64118	Q64118 meriones un
36	184.5	16.1	418	11	Q91XB0	Q91XB0 mus musculus
37	184.5	16.1	418	11	Q8VCH3	Q8VCH3 mus musculus
38	184	16.1	418	5	Q27085	Q27085 tachyples
39	183.5	16.0	413	11	Q63969	Q63969 mus baxicol
40	182.5	15.9	380	11	Q9D695	Q9D695 mus musculus
41	182.5	15.9	418	11	Q91W80	Q91W80 mus musculus
42	182	15.9	396	6	O00394	O00394 cercopithec
43	182	15.9	413	11	O54761	O54761 spermophilu
44	180.5	15.8	433	13	Q9Y1B8	Q9Y1B8 xenopus lae
45	179	15.6	413	6	Q28666	Q28666 oryctolagus

#### ALIGNMENTS

RESULT 1

ID	O08804	PRELIMINARY;	PRT;	377 AA.
AC	O08804;			
DT	01-JUL-1997 (TRENBLrel. 04, Created)			
DT	01-DEC-2001 (TRENBLrel. 19, Last sequence update)			
DT	01-JUN-2002 (TRENBLrel. 21, Last annotation update)			
DE	Serine proteinase inhibitor NK13.			
GN	SP112.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_faxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CS7BL;			
RX	MEDLINE=97326124; PubMed=9182575;			
RA	Sun J., Ooms L., Bird C.H., Sutton V.R., Trapani J.A., Bird P.I.;			
RT	"A new family of 10 murine ovalbumin serpins includes two homologs of			
RT	proteinase inhibitor 8 and two homologs of the granzyme B inhibitor			
RT	(proteinase inhibitor 9).";			
RL	J. Biol. Chem. 272:15434-15441 (1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CS7BL;			
RA	Kaiserman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,			
RA	McKeone R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,			
RA	Ragoussis J., Bird P.I.;			
RT	"Comparison of human chromosome 6p25 with murine chromosome 13 reveals			
RT	a greatly expanded Ov-serpin gene repertoire in the mouse.";			
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.			
CC	-1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.			
DR	EMBL; U96707; AAB57819.2; -			
DR	MGI; MGI:894688; Sp112.			
DR	InterPro; IPR000215; Serpin.			
DR	Pfam; PF00079; serpin; 1.			
DR	SMART; SM00093; SERPIN; 1.			
DR	PROSITE; PS00284; SERPIN; 1.			
KW	Serpin.			

SEQUENCE 377 AA; 42535 MW; 9660E8F3A41BE47A CRC64;

Query Match 33.0%; Score 377.5; DB 11; Length 377;  
Best Local Similarity 50.7%; Pred. No. 2.5e-29;  
Matches 75; Conservative 29; Mismatches 41; Indels 3; Gaps 2;

DQ  
41 KOLIFFPMISISALAMVFMGAKNTAAOMSQAALCFSGKIGEDG-DIHRGFQSLLVAINRT 99  
: : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 25 RNVLSPISVSALAMVFMGAGTTTASQAQALSLDCKSGKGGRDVHQFSQLLTETNKT 84

DQ  
100 DTEYLVLRITANGLFGEKSDFLGTFTDSCGKFYAQTIKOLDVFNDTPKSTTRVNSWADKT 159  
: : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 85 GTQYVLRITANRLEFGKTFDLASFKDSCRKYEAEMEELDFKGATEQSRQHINAWVKIT 144

DQ  
160 --KWKIIQTSLSHLEEPIASSCYCK 185  
: : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 145 EDKITELLSSGSVNSNTPLVLNAVIFYK 172

RESULT 2

OQ  
08VHQO; PRELIMINARY; PRT; 385 AA.

AC  
01-WAR-2002 (TREMBLrel. 20, Created)  
DT  
01-WAR-2002 (TREMBLrel. 20, Last sequence update)  
DDT  
01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE SPI3J2.  
GN SP13J2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B6D2 F1/J;  
RA Kaiserman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,  
RA McKeone R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,  
RA Rogussis J., Bird P.I.;  
RT "Comparison of Human Chromosome 6p25 with Murine Chromosome 13 Reveals  
RL a Greatly Expanded Ov-Serpin Gene Repertoire in the Mouse."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.  
DR ENBL; AF425084; AAL65910.1; -  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF00079; serpin; 1.  
DR SMART; SM00093; SERPIN; 1.  
DR PROSITE; PS00284; SERPIN; UNKNOWN 1.  
SQ SEQUENCE 385 AA; 43786 MW; 0820A8A26AE8B485 CRC64;

Query Match 30.3%; Score 347.5; DB 11; Length 385;  
Best Local Similarity 48.0%; Pred. No. 2.5e-26;  
Matches 71; Conservative 25; Mismatches 49; Indels 3; Gaps 2;

DQ  
41 KOLIFFPMISISALAMVFMGAKNTAAOMSQAALCFSKI-GGEDGIHRGFQSLLVAINRT 99  
: : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 25 KNVFLSPIISSIALVMVGLRGKGGTAITQITQALSCKSSSEDGDVGHFQOLLSEVNKT 84

DQ  
100 DTEYLVLRITANGLFGEKSDFLGTFTDSCGKFYAQTIKOLDVFNDTPKSTTRVNSWADKT 159  
: : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 85 GTQYSLKAANRLFGEKTFDLASFKDSCHKFYEAEMEELDFKGATEQSRQHINTVAVKT 144

DQ  
160 --KWKIIQTSLSHLEEPIASSCYCK 185  
: : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 145 EDKIKELLSPGTIHGNTPLILNAVIFYK 172

RESULT 3

OQ  
008800 PRELIMINARY; PRT; 374 AA.

ID O08800

AC O08800;

DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DDT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE SWATASE inhibitor NK10 (serine protease inhibitor 8).

SQ SEQUENCE 369 AA; 42287 MW; CSD7A2426378DE3E CRC64;  
 Query Match 23.8%; Score 272; DB 4; Length 369;  
 Best Local Similarity 38.6%; Pred. No. 8.4e-19;  
 Matches 54; Conservative 27; Mismatches 47; Indels 12; Gaps 1;  
 QY 32 FEKARGKQLKLIFFPMSISSALAMVFMGAKGNTAAQMSQALCFKSKIG----- 79  
 DB 16 FOQFRKKNENIFSPISITSSALGMVLLGAKDNTAQISKVLHPDQVTENTTEKAAHYV 75  
 QY 80 GEDGDIHRGFSQSLVAINTDEYVLTANGLCFGEKSYDFLTGFTDSCGKGYQATIKOLD 139  
 DB 76 DRSGNVHQFQKLTENKSTDAELKANKLFGEKTYQLQEYLDIAIKKFYQTSVESTD 135  
 QY 140 FVNDTEKSTRVNSWADKT 159  
 DB 136 FANAPESRKKINSWVESQT 155  
 RESULT 5  
 ID Q9BYF8 PRELIMINARY; PRT; 338 AA.  
 AC Q9BYF8;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE SCCALb.  
 GN SCCAL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Suminami Y., Kishi F., Murakami A., Sakaguchi Y., Kato H.;  
 RT "Novel Forms of SCC Antigen Transcripts Produced by Alternative  
 RT Splicing."  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 DR EMBL; AB046399; BAB40772.1; -.  
 DR HSSP; P01008; IATH.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 DR Serpin.  
 KW Serpin.  
 SQ SEQUENCE 338 AA; 38519 MW; 5AB9CBF480CIABDF CRC64;  
 Query Match 23.2%; Score 266; DB 4; Length 338;  
 Best Local Similarity 37.9%; Pred. No. 3e-18;  
 Matches 53; Conservative 28; Mismatches 47; Indels 12; Gaps 1;  
 QY 32 FEKARGKQLKLIFFPMSISSALAMVFMGAKGNTAAQMSQALCFKSKIG----- 79  
 DB 16 FOQFRKKNENIFSPISITSSALGMVLLGAKDNTAQIKKVLHPDQVTENTTEKAAHYV 75  
 QY 80 GEDGDIHRGFSQSLVAINTDEYVLTANGLCFGEKSYDFLTGFTDSCGKGYQATIKOLD 139  
 DB 76 DRSGNVHQFQKLTENKSTDAELKANKLFGEKTYQLQEYLDIAIKKFYQTSVESTD 135  
 QY 140 FVNDTEKSTRVNSWADKT 159  
 DB 136 FANAPESRKKINSWVESQT 155  
 RESULT 6  
 ID O08797 PRELIMINARY; PRT; 374 AA.  
 AC O08797;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Serine proteinase inhibitor 6.

GN SPI6 OR SPI6.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sun J., Ooms L., Bird C.H., Sutton V.R., Trapani J.A., Bird P.I.;  
 RT "A new family of ten murine ovalbumin serpins includes two homologs of  
 RT proteinase inhibitor 8 and two homologs of the granzyme B inhibitor  
 RT (proteinase inhibitor 9)."  
 RL J. Biol. Chem. 0:0-0(1997).  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 DR EMBL; U96700; AAB57812.1; -.  
 DR HSSP; P05619; IHLE.  
 DR MGD; MGI:106603; Spi6.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 DR Serpin.  
 KW Serpin.  
 SQ SEQUENCE 374 AA; 42259 MW; 76D9ED10D7875281 CRC64;  
 Query Match 22.6%; Score 259; DB 11; Length 374;  
 Best Local Similarity 46.7%; Pred. No. 1.7e-17;  
 Matches 56; Conservative 23; Mismatches 37; Indels 4; Gaps 2;  
 QY 41 KOLIFPMSISSALAMVFMGAKGNTAAQMSQALCFKSKIGGEDIHGFQSLVAINTRD 100  
 DB 26 KNCVSPASISSALAMVLLGAKGTAVQISQALGLNK--BEG-IHQGFQSLVAINTKPD 81  
 QY 101 TEYVLTANGLFGEKSYDFLTGFTDSCGKGYQATIKOLDVNDTEKSTRVNSWADTK 160  
 DB 82 RYSLRVANRLFADTKCEVLTQTFKESLHFYDSEMQLSFAEAEVSRQHTVWSKQTE 141  
 RESULT 7  
 ID Q8VHP7 PRELIMINARY; PRT; 382 AA.  
 AC Q8VHP7;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE EIB.  
 GN SERPINB18.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C; TISSUE=LUNG;  
 RA Kaiserman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,  
 RA McKeone R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,  
 RA Ragousis J., Bird P.I.;  
 RT "Comparison of human chromosome 6p25 with murine chromosome 13 reveals  
 RT a greatly expanded Ov-serpin gene repertoire in the mouse."  
 RL Genomics 0:0-0(2002).  
 DR EMBL; AF426025; AAL57487.1; -.  
 DR InterPro; IPR002016; Peroxidase.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00435; PEROXIDASE 1; UNKNOWN\_1.  
 DR PROSITE; PS00284; SERPIN; UNKNOWN\_1.  
 DR Serpin.  
 KW Serpin.  
 SQ SEQUENCE 382 AA; 42887 MW; 35CB86ADF677C8DB CRC64;  
 Query Match 22.0%; Score 251.5; DB 11; Length 382;  
 Best Local Similarity 36.8%; Pred. No. 9.8e-17;  
 Matches 57; Conservative 31; Mismatches 50; Indels 17; Gaps 3;  
 QY 43 LIFPMSISSALAMVFMGAKGNTAAQMSQALCFKSKIGGEDIHGFQSLVAINTDTE 102

Wed May 28 14:44:11 2003

```

Db 28 IFFSPFSSISLAWFLGKAGSTRAQLSKTLHFDV-----EDHSCFOSLTAEVSKLGAS 83
Qy 103 YVLRTRANGLFGEKSYDFLTGFTDSCGKYQATIKQLDFVNDTEKSTTRVNSWVADTKAW 162
Db 84 HTLKLNRLYGKTYNLFPEFLASTQKYSADLAADVDFHASEDARKEINQWVGQTEG- 142
Qy 163 KIIQTSLSHLEPPGIASSCYKACLSOPLLVHSI 197
Db 143 -----KIPPELLAGVDS-----MTKLVLVNAI 165

RESULT 8
O0806 PRELIMINARY; PRT; 377 AA.
AC O0806
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Serine proteinase inhibitor NK26.
GN SPI14.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -|- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AK005394; BAB2398.1; --
DR HSSP; P05120; 1BY7.
DR MGD; MGI:1919260; 1600002F03rik.
DR InterPro; IPR000240; Maspin.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR PRINTS; PR00676; MASPIN.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin.
SQ SEQUENCE 377 AA; 43025 MW; CD162840210BE40D CRC64;

Query Match 21.7%; Score 248.5; DB 11; Length 377;
Best Local Similarity 45.0%; Pred. No. 1.9e-16;
Matches 54; Conservative 22; Mismatches 39; Indels 5; Gaps 2;

Qy 41 KOLIFFPMSISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFSQSLVAINRTD 100
Db 26 KNCVSPMSISSALAMVLLGAGDTAVQICQALHLN----PDEDVHGFQQLLHNLKN 81
Qy 101 TE-YVLRTRANGLFGEKSYDFLTGFTDSCGKYQATIKQLDFVNDTEKSTTRVNSWVADTK 159
Db 82 NQYCLTWANRLFVENTCELLPTFKSKLKYHSEIQLSFAEAESRQHINMWVSKQT 141

RESULT 10
Q8VHQ1 PRELIMINARY; PRT; 377 AA.
ID Q8VHQ1
AC Q8VHQ1;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NK21L1.
GN NK21L1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

[illegible]

Wed May 28 14:44:11 2003

SQ SEQUENCE 379 AA; 42575 MW; F492FLEEB8880B1 CRC64;  
 Query Match 21.2%; Score 243; DB 11; Length 379;  
 Best Local Similarity 34.5%; Pred. No. 6.9e-16; Indels 24; Gaps 4;  
 Matches 58; Conservative 32; Mismatches 54;

QY 43 LIFPMSISSALAMVPMGAKNTAAQMSQALCFSGKIGDGDGDIHRGFSQLLVAINTDTE 102  
 Db 28 IFFSPFSSISSALAMVILGAKGSTAAQLSKTFHDSV-----EDIHSRFQSLNAEVSXRGAS 83  
 QY 103 YVLRRTANGLFGEKSYDFLTGFTDSCGKYQATIKQLDFVNDTEKSTTRVNSWADTKAW 162  
 Db 84 HTLKLNLRYGKTYNLFPEYLASTQKMGADLAPVDFLHASEDARKEINQWVGQTEG- 142  
 QY 163 KIIQTSLSHLEBPGFIASSCYKACLSQPLLVHSIPKCNVPTVPHGMW 210  
 Db 143 -----KIPKLLSVGVVNS-----MTKLVLVNAI-----YFKGMW 171

RESULT 13  
 Q9D154 PRELIMINARY; PRT; 379 AA.  
 ID Q9D154  
 AC Q9D154  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE 119005M04RIK protein (RIKEN cDNA 119005M04 gene) (EIT).  
 GN SERPINB1 OR 119005M04RIK OR SERPINB1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Kamiya M., Lee N.H.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Storch K.-F.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=SPLEEN;  
 RC Kaiserman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,  
 RC McKeone R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,  
 RA Ragoussis J., Bird P.I.;  
 RA "A new family of 10 murine ovalbumin serpins includes two homologs of  
 RT a greatly expanded Ov-serpin gene repertoire in the mouse."  
 RL Genomics 0:0-0(2002).  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 DR EMBL; AK003930; BAB23079.1; -;  
 DR EMBL; BC011140; AAL11140.1; -;  
 DR EMBL; AF426024; AAL57486.1; -;  
 DR HSSP; P05619; 1HLE.  
 DR MGP; MGI:1913472; Serpinb1a.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 W Serpin.

SQ SEQUENCE 379 AA; 42575 MW; F492FLEEB8880B1 CRC64;  
 Query Match 21.2%; Score 243; DB 11; Length 379;  
 Best Local Similarity 34.5%; Pred. No. 6.9e-16; Indels 24; Gaps 4;  
 Matches 58; Conservative 32; Mismatches 54;

QY 43 LIFPMSISSALAMVPMGAKNTAAQMSQALCFSGKIGDGDGDIHRGFSQLLVAINTDTE 102  
 Db 28 IFFSPFSSISSALAMVILGAKGSTAAQLSKTFHDSV-----EDIHSRFQSLNAEVSXRGAS 83  
 QY 103 YVLRRTANGLFGEKSYDFLTGFTDSCGKYQATIKQLDFVNDTEKSTTRVNSWADTKAW 162  
 Db 84 HTLKLNLRYGKTYNLFPEYLASTQKMGADLAPVDFLHASEDARKEINQWVGQTEG- 142  
 QY 163 KIIQTSLSHLEBPGFIASSCYKACLSQPLLVHSIPKCNVPTVPHGMW 210  
 Db 143 -----KIPKLLSVGVVNS-----MTKLVLVNAI-----YFKGMW 171

RESULT 14  
 Q9DAV6 PRELIMINARY; PRT; 377 AA.  
 ID Q9DAV6  
 AC Q9DAV6; O08802;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE 1600019A2IRIK protein (Serine proteinase inhibitor R86).  
 GN 1600019A2IRIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=PLACENTA;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Kamiya M., Lee N.H.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Storch K.-F.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=WHOLE EMBRYO;  
 RC MEDLINE=97326124; PubMed=9182575;  
 RA Sun J., Ooms L., Bird C.H., Sutton V.R., Trapani J.A., Bird P.I.;  
 RA "A new family of 10 murine ovalbumin serpins includes two homologs of  
 RT proteinase inhibitor 8 and two homologs of the granzyme B inhibitor  
 RL (proteinase inhibitor 9)."  
 RL J. Biol. Chem. 272:15434-15441(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=WHOLE EMBRYO;  
 RC Kaiserman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,  
 RA McKeone R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,  
 RA Ragoussis J., Bird P.I.;  
 RA "Comparison of human chorion gene repertoire in the mouse."  
 RT a greatly expanded Ov-serpin gene repertoire in the mouse."



Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.

CC EMBL; AK005491; BAB24076.1; --  
 DR EMBL; U96705; AAB57817.2; --  
 DR HSP; P05619; LHLE  
 DR MGD; MGI:1919461; 1600019A21Rik.  
 DR InterPro; IPR000240; Maepin.  
 DR Pfam; PF000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR PRINTS; PRO0676; MASPIN.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Serpin.

SQ SEQUENCE 377 AA; 42733 MW; B6DCAE1ADE932F4 CRC64;

Query Match 21.0%; Score 240; DB 11; Length 377;

Best Local Similarity 43.3%; Pred No. 1.4e-15;

Matches 52; Conservative 24; Mismatches 40; Indels 4; Gaps 1;

QY 41 KQLIFFPMISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFQSLVAINRTD 100

Db 26 KNVCFSPVSISSALAMVLLGAEQTAVQISQALGLKKEG----IHQFLKLLRLINKED 81

QY 101 TEVLTANGLFGEKSYDLTGTDCGKFKYQATIKQLDFVNDTEKSTTRVNSWVADTKY 160

Db 82 RKYSLLVANRLFADKTCVLTQTFKSCFRFDSEMEQVNFKAARQCIINTVWSQTE 141

#### RESULT 15

Q9D7S8  
 ID Q9D7S8 PRELIMINARY; PRT; 379 AA.  
 AC Q9D7S8;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE 1190005M04Rik protein.  
 GN SERPINB1A OR 1190005M04RIK.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=STOMACH;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itch M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Mizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Aito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bareh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayaebizaki Y.,

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.

DR EMBL; AK008914; BAB25964.1; --

DR HSP; P05619; LHLE

DR MGD; MGI:1913472; Serpinb1a.

DR InterPro; IPR000215; Serpin.

DR Pfam; PF00079; serpin; 1.

DR SMART; SM00093; SERPIN; 1.

DR PROSITE; PS00284; SERPIN; 1.

KW Serpin.  
 SQ SEQUENCE 379 AA; 42580 MW; 193728585077AB4B CRC64;  
 Query Match 20.7%; Score 237; DB 11; Length 379;  
 Best Local Similarity 33.9%; Pred. No. 2.7e-15;  
 Matches 57; Conservative 32; Mismatches 55; Indels 24; Gaps 4;  
 QY 43 LIFFPMISSALAMVFMGAKGNTAAQMSQALCFSGKIGGEDGDIHRGFQSLVAINRTDTE 102  
 Db 28 IFFSPFSSISALAMVILGAKGTAQAQLSKTRHFDV-----EDIHSRFQSQNAEVSQRGAS 83  
 QY 103 YVLTANGLFGEKSYDLTGTDCGKFKYQATIKQLDFVNDTEKSTTRVNSWVADTKAW 162  
 Db 84 HTLKLANRLYGEKTYNLFPEYLASTQKMYGADLAPVDFLHASEDARKEINQWVKQTEG- 142  
 QY 163 KIIQTSLSHLSEEPGIASSCYCKACLSQLLVHSIPKCNPSVTPHGMW 210  
 Db 143 -----KIPELLISVGVDVDS-----MTKLVLVNAI-----YPKGMW 171

Search completed: May 20, 2003, 10:25:54  
 Job time : 88 secs

